Microscopic Hamiltonian Systems and their Effective Description

Sutton, Daniel

Award date:
2013

Awarding institution:
University of Bath

Microscopic Hamiltonian Systems and their Effective Description

submitted by

Daniel Colin Sutton

for the degree of Doctor of Philosophy

of the

University of Bath

Department of Mathematical Sciences

May 2013

Contents

Acknowledgements

I would like to express my deep gratitude to Hartmut Schwetlick and Johannes Zimmer, my research supervisors, for their patient guidance, enthusiastic encouragement and useful critiques of this research work. Working with them I have learned a lot, both about life and mathematics. One of the more prominent things I have learned, despite being a native English speaker, is that I cannot write or speak English very well. I'm sure this will become apparent very soon.

I would also like to thank Jey Sivaloganathan for sparking my interest in applied analysis. I would also like to thank all of the staff members of the mathematics department at Bath, between the years of 2006-2013, for teaching me mathematics and their interesting discussions. Further to this, I would also like to thank my tutees from 2008-2012, for being a pleasure to work with.

In addition I would also like to thank Adam Boden, Joe Collins, Shane Cooper, Hong Duong, Andrea Fernandez, Horacio González, Ellie Harrison, Vaios Lachos, Christian Mönch, Aretha Teckentrup and Alex Watson, all for providing both support and useful discussions. I would also like to extend my thanks to Mary Baines, Owen Cranshaw, Sarah Hardy, Ann Linfield, Carole Negre, Ade Olatokun, Mark Willis and Eric Wing, for their help and support.

Finally, I wish to thank my family and friends for their support and encouragement throughout my study.

List of Figures

List of Procedures

Notation

Topology and Metric Spaces

$\text{Image}(u)$	Image of the mapping u
$\text{cl}(A)$	Closure of the set A
$\text{int}(A)$	Interior of the set A
∂A	Topological boundary of A
$\text{diam}(A)$	The diameter of the set A
$B_\rho(x)$	Metric ball of radius ρ and centre x
$A \subset\subset B$	A is a compact subset of B
$\|x\|$	The Euclidean norm

Function Spaces

BUC	Bounded uniformly continuous functions
L^p	All $p-$integrable functions
$W^{m,p}$	The Sobolev space of functions with m, $p-$integrable, weak derivatives
$\mathcal{A}(\xi_1, \xi_2)$	All Lipschitz curves joining ξ_1 to ξ_2.

Chapter 1

Overview

The universal theme connecting the topics of this thesis is how microscopic details can help us understand the macroscopic behaviour of Hamiltonian systems. We will focus on three aspects of Hamiltonian dynamics which we now describe, this is not intended to be a comprehensive description, we provide a comprehensive introduction to each aspect in the corresponding chapters.

The second chapter of this thesis focuses on a model for phase transitions in an atomistic model of a solid. A reason for studying atomistic models for elastic solids is that on the macroscopic level there exist many solutions representing dynamic phase transitions. Ultimately the question arises, which solution do we expect to see in reality? One selection criterion is the kinetic relation, that is, a relationship between the configurational force and the velocity of the phase boundary. As such, a new problem arises, how does one determine the kinetic relation? While there are many approaches to determining such relations, one approach is to derive the relation from first principles, this is possible through the study of the microscopic lattice model that corresponds to the macroscopic material. This was done in [SZ09b] for a particular potential. The equations governing the motion in the case of a nonconvex interaction potential have only recently started being analysed in a rigourous fashion. The equation that describes the motion of a one dimensional FPU lattice, the atomistic model of a one dimensional elastic solid, is

$$\frac{\mathrm{d}^2 x_n}{\mathrm{d}t^2} = V'(x_{n+1} - x_n) - V'(x_n - x_{n-1}), \quad n \in \mathbb{Z}.$$

The particular potential that we study is

$$V(x) = \frac{1}{2} \min\left\{(x+1)^2, (x-1)^2\right\},$$

a model nonconvex potential. It proves to be an interesting problem to study the exis-

tence of solutions to this problem, and for more general potentials V. In [SZ09b] they find that a class of traveling wave solutions exist for a range of velocities just below the speed of sound. The key result of the first chapter is that we show that for substantially lower velocities that these travelling wave solutions no longer exist. Such a dichotomy of behaviour has been previously observed experimentally in [FS40].

The third chapter studies the homogenisation of metric functionals. The aim is to study the Γ-limit of a sequence of rapidly oscillating metric functionals. Specifically, functionals of the form

$$F_\varepsilon(u) = \int_0^1 a\left(\frac{u(\tau)}{\varepsilon}\right) \|u'(\tau)\| \mathrm{d}\tau.$$

While computing the minimal curves in such functionals may prove to be a computationally expensive task, the process of homogenisation allows us to solve a structurally simpler problem, which is a suitable approximation of the original minimal curve. The purpose of the chapter is to contribute further to the theory of the homogenisation of such functions, and how it applies to the homogenisation of Hamiltonian dynamics. We will first examine the results of [BPF01] where they study a general relationship between convergence of metrics and the Γ-convergence of metric functionals. In particular, we will revisit these ideas in the context of the homogenisation of rapidly oscillating metrics to find additional structure. The next problem we examine is the homogenisation of a rapidly oscillating metric taking values in either $\{1, \infty\}$ or $\{1, \beta\varepsilon^{-p}\}$, where $\beta \in \mathbb{R}$ and ε is the size of the period cell. The existing theory in the literature fails to apply to such metrics. The results of the analysis of the metrics, in the context of Γ-convergence, exhibit a new type of behaviour. We observe that the length functionals will always Γ-converge, however, the problem of determining the Γ-limit for the boundary value problem behaves differently. Typically, in the case of uniformly bounded functionals, it suffices to determine the unconstrained Γ-limit and use this to show that the Γ-limit exists for the boundary value problem [BD98, Chapter 11]. What we observe is that for a sequence of metrics that fail to be uniformly bounded, the existence of the Γ-limit depends on the value of p. Should $p < 1$ then the Γ-limit exists, should $p \geq 1$ or the metric takes values in $\{1, \infty\}$ then the Γ-limit fails to exist. The next metric homogenisation problem we study is a particular example, where we can determine the Γ-limit explicitly. For the example we take

$$a_\rho(x) := \begin{cases} \beta, & \text{if } x \in \frac{1}{2}(1-\rho, 1+\rho)^2, \\ 1, & \text{if } x \in [0,1]^2 \setminus \frac{1}{2}(1-\rho, 1+\rho)^2, \end{cases}$$

where $\beta > 2$, $\rho \in (0, 1)$. The interesting, and previously unobserved, feature of this example is that the limit metric is piecewise affine, whose level sets form $\infty-$gons. This

has clear implications for those wishing to develop numerical methods to solve such homogenisation problems. The other interesting feature of this example is that the metric depends on a parameter ρ, and therefore for the first time, we are able to study the effect of changing the microscopic properties of the metric on the homogenised limit. Finally, we study the homogenisation of the Hamiltonian boundary value problem, using Maupertuis' metric formulation of Hamiltonian dynamics. Our conclusion is that if one wishes to determine the effective behaviour of trajectories in a rapidly oscillating potential then both our approach and the approach of [LPV88] provide identical dynamical information.

The fourth, and final, chapter studies the numerical computation of geodesics to solve the Hamiltonian boundary value problem. Using Maupertuis' metric formulation of Newton's second law we study various curve shortening procedures to find numerical solutions of

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = -\nabla V(x),$$

when posed as a boundary value problem. Our reasoning for studying this particular problem is to calculate long time transitions in molecular systems. In this work we will demonstrate the calculation of a particular molecular transition in the change of conformation for the butane molecule at low energies. In [SZ09a, SZb, SZa] they have developed a consistent approach to the computation of extended minimal curves for Riemannian metrics. The aim of the work is to modify the existing curve shortening algorithms to run faster in commonly encountered situations, using a predictor-corrector method. We will demonstrate, using an example problem, that the proposed modifications do indeed lead to an improvement. In addition to this we will also explore an implementation of the modified algorithm in a parallel architecture.

There will be three chapters, each of which covering an aspect of Hamiltonian dynamics. Each chapter contains its own introduction, complete with a detailed discussion of the relevant background literature. The contributions to the literature made by this thesis are detailed in each chapter's introduction. For the final chapter there are also associated code listings, included in the appendix.

Publications

The result of chapter 2 has been published as "Nonexistence of Slow Heteroclinic Travelling Waves for a Bistable Hamiltonian Lattice Model" in the Journal of Nonlinear Science (**22**, 2012, 917–934). The results of section 3.4 have been submitted to the Journal of Convex Analysis, provisionally entitled "The Finsler Metric Obtained as the Γ-limit of a Generalised Manhattan Metric".

Chapter 2

Nonexistence of Travelling Waves in a Bistable Lattice Model

2.1 Introduction

In this chapter we study the nonexistence of travelling waves for a bistable Hamiltonian lattice model. The purpose of this study is to show that travelling waves, which serves as a model for the motion of a phase boundary, fail to exist for low wave speeds. We perform this study using a Fermi-Pasta-Ulam (FPU) chain to model the material atomistically. The FPU chain we consider here is a one-dimensional, bi-infinite chain of identical point unit masses, representing the atoms, joined to their nearest neighbours with nonlinear springs. The springs typically have a nonconvex stored energy potential with different wells representing the different stable phases. We consider the simplest multiphase material, that is with two distinct stable phases. For our analysis consider a particular stored energy potential that is piecewise quadratic which is defined in (2.4). This model with piecewise quadratic interactions was studied analytically and numerically in [BCS00a, BCS00b].

Let $u_j(t) \in \mathbb{R}$ be the displacement of the jth atom with respect to the uniform reference configuration $\mathbb{Z}$ at time $t \in \mathbb{R}$. Denoting the potential function as $V\colon \mathbb{R} \to \mathbb{R}$ and assuming that the evolution of the dynamics is governed by Newton's second law, one finds that the equation of motion is

$$\ddot{u}_j(t) = V'(u_{j+1}(t) - u_j(t)) - V'(u_j(t) - u_{j-1}(t)), \qquad j \in \mathbb{Z}. \tag{2.1}$$

Some parts of this chapter have appeared in [SSZ12].

A solution of (2.1) is a *travelling wave* if it has the form

$$u_j(t) = u(j - ct), \qquad j \in \mathbb{Z}, \tag{2.2}$$

where the constant c is the wave speed. We say that a travelling wave solution represents a *phase transition* in the material if it has strains in both wells of the potential. Furthermore, a travelling wave representing a phase transition is *heteroclinic* if it asymptotically belongs to different wells. Such phase transitional travelling waves were first studied using Fourier analysis for a FPU chain with piecewise quadratic interaction potential in [TV05]. The approach in [TV05] is to assume that the solution of (2.1) takes a particular form, that is, the profile of the travelling wave only crosses between wells once. With this assumption it is possible to reduce the nonlinear equation (2.1) to a linear inhomogeneous equation, see section 2.3. Consequently, the latter equation is easier to analyse using Fourier methods. In [TV05] they express the solution as a formal sum, the problem being that it is difficult to verify whether the sum satisfies the single transition assumption. In [SZ09b] an alternative framework to address the existence of subsonic phase transition waves very close to the speed of sound is proposed. The speed of sound of the material is denoted by c_0 and is determined by calculating $\sqrt{V''}$. The idea is that the solution of the linear inhomogeneous equation as the sum of a given function, the profile, and a remainder term, the corrector. The profile is chosen to ensure that the corrector is square integrable and the Fourier transform of the corrector is known. Therefore, it is possible to make detailed estimates on the values of the corrector function, with the values of the profile function known it could be deduced that the resulting sum satisfies the single transition property. Here we show that this framework, although used to prove existence, can be adapted to prove a contrary proposition, the nonexistence of single transition waves for a slow wave speed regime.

The question of what happens at subsonic wave speeds significantly lower than the speed of sound has, to the best of the author's knowledge, not been addressed in an analytical framework before. It has been conjectured by Peyrard and Kruskal [PK84] that travelling waves with low constant wave speeds do not exist for the related Frenkel-Kontorova model on finite domains, based on numerical computations. Here we show this conjecture is true for the bi-infinite FPU chain as there is no travelling wave joining bounded strains in the different wells of the bilinear potential for wave speeds significantly lower than the speed of sound. Consequently this means that at low subsonic wave speeds there are no phase transitional solutions to the lattice differential equation (2.1) that makes a single transition between the potential wells. Remarkably, the methods are rather similar to those used to show the opposite result, namely the the existence

of travelling waves for very fast subsonic waves [SZ09b]. Our result indicates that the motions at the low wave speeds considered here may be less coherent than these with speeds close to the speed of sound. It may be possible that there are travelling wave solutions with multiple interfaces, or solutions that are not of travelling wave type. In conjunction with [SZ09b], the result presented here describes a dichotomy: coherent single-interface travelling waves exist for high subsonic velocities but not for low velocity. Such a dichotomy between fast and slow martensitic transformations has been observed experimentally by Förster and Scheil [FS40] in the 1940's. In [FS40] they studied the behaviour of transitions in steel that is suddenly cooled. They observed that there exists two types of transitions. The first of these transformations is a high speed ordered transition, characterised by a 'clapping' sound. We believe that such transitions are described by the waves in [SZ09b]. The other type of transition is a slow and disordered transition. It would be of interest to understand how the latter transformation manifests in our atomistic model.

The existence results of [SZ09b] have been extended further in [HMSZ13] to include a small spinodal region in the potential. The result was obtained by including a sufficiently small perturbation of the potential and applying a fixed point argument to the resulting problem in relation to the solution of [SZ09b]. Further lattice models have been explored in the context of formal series in [Vai10, VK12, VHRT98].

2.2 Mathematical Description

We consider a one-dimensional chain of atoms $\{q_j\}_{j\in\mathbb{Z}} \subset \mathbb{R}$ whose deformations are given as $u_j \colon \mathbb{R} \to \mathbb{R}$. We have made the assumption that the dynamics can be described by Newton's second law and that the equations of motion are given by (2.1).

The motion of the phase boundary can be modelled as a travelling wave with strains in both wells of the potential. With the ansatz (2.2) the equations of motion (2.1) reduce to a single equation

$$c^2 u''(x) = V'(u(x+1) - u(x)) - V'(u(x) - u(x-1)). \tag{2.3}$$

For the analysis of phase transitions in lattice models it is beneficial to reformulate equation (2.3) in terms of the *discrete strain*. We define the discrete strain as $\varepsilon(x) := u(x) - u(x-1)$ and specify the potential as a function of ε. In this study we consider the potential previously analysed in [BCS00a, BCS00b, SZ09b, TV05],

$$V(\varepsilon) := \frac{1}{2} \min\left\{(\varepsilon+1)^2, (\varepsilon-1)^2\right\}. \tag{2.4}$$

So there are two wells joined at 0 by a corner and the speed of sound is unity. Having wells at ± 1 is immaterial however it is possible to rescale and translate the potential, as demonstrated by Schwetlick and Zimmer in [SZ07], so that the wells are located at 0 and at a small positive strain. Furthermore, we define the discrete Laplacian to be

$$\Delta_1 f(x) := f(x+1) - 2f(x) + f(x-1).$$

Equation (2.3) can be now reformulated as the *discrete strain equation*

$$c^2 \varepsilon''(x) = \Delta_1 V'(\varepsilon(x)), \tag{2.5}$$

where we take $V'(0) = 0$. Given the explicit form of the potential (2.4) it is easy to check that (2.5) becomes

$$c^2 \varepsilon''(x) = \Delta_1 \varepsilon(x) - 2\Delta_1 H(\varepsilon(x)), \tag{2.6}$$

where

$$H(x) := \begin{cases} 0 & \text{if } x < 0, \\ \frac{1}{2} & \text{if } x = 0, \\ 1 & \text{if } x > 0. \end{cases} \tag{2.7}$$

Defining the linear operator $L_c := c^2\partial^2 - \Delta_1$ we rewrite (2.6) as the following nonlinear advance-delay differential equation

$$L_c \varepsilon(x) = -2\Delta_1 H(\varepsilon(x)). \tag{2.8}$$

We say that a travelling wave satisfies the *sign condition* or has a *single transition* if it satisfies the property

$$x \cdot \varepsilon(x) > 0 \text{ for every } x \neq 0. \tag{SC}$$

Condition (SC) is central to this chapter as it implies that there is exactly one transition between the potential wells, located at the origin in the moving frame coordinates.

The aim of this chapter is to demonstrate that there exists a range of values for c, whose absolute values are much less than unity, such that there are no single-transition heteroclinic travelling wave solutions to (2.8).

2.3 Fourier Analysis and the Dispersion Relation

The Fourier transform of an $L^1(\mathbb{R})$ function $u\colon \mathbb{R} \to \mathbb{R}$ is

$$\mathscr{F}[u](\kappa) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u(x) \exp(-i\kappa x) \mathrm{d}x, \ \kappa \in \mathbb{R}$$

where this exists. The Fourier sine transform of u is

$$\mathscr{F}_s[u](\kappa) := \sqrt{\frac{2}{\pi}} \int_0^{\infty} \sin(\kappa x) u(x) \mathrm{d}x. \tag{2.9}$$

Note that the relation $\mathscr{F}[u] = -i\mathscr{F}_s[u]$ holds when u is an odd function.

We define the dispersion relation to be the symbol of the linear operator L_c, which is defined by

$$\mathscr{F}[L_c\varepsilon](\kappa) = D(\kappa, c)\mathscr{F}[\varepsilon](\kappa),$$

where

$$D(\kappa, c) := -c^2\kappa^2 + 4\sin^2\left(\tfrac{1}{2}\kappa\right) \tag{2.10}$$

is the *dispersion function*. The dispersion relation is given by $D(\kappa, c) = 0$. It proves convenient to define the function

$$d(\kappa) := \begin{cases} \left(\frac{\sin\left(\frac{1}{2}\kappa\right)}{\frac{1}{2}\kappa}\right)^2 & \text{if } \kappa \neq 0, \\ 1 & \text{if } \kappa = 0 \end{cases} \tag{2.11}$$

so that we can rewrite the dispersion relation as $D(\kappa, c) = (d(\kappa) - c^2)\kappa^2$. As a consequence κ is a zero of the dispersion relation if and only if $d(\kappa) = c^2$ or $\kappa = 0$ for any $c \in \mathbb{R}$.

In this chapter we consider values of c for which the equation $d(\kappa) = c^2$ has precisely three roots although in principal these arguments hold for c corresponding to a higher odd number of roots. The case with $c = 0.016$, corresponding to 5 distinct roots, is considered in Section 2.6. These situations has been studied numerically by Slepyan *et al.* in [SCC05]. Instead of specifying the wave speed directly we prescribe a root of the dispersion relation. This in turn defines the wave speed and the other roots. Let $\hat{\kappa}$ be the value of κ corresponding to the unique maximum of d on $[2\pi, 4\pi]$. We will use numerical approximations of the value of $\hat{\kappa}$ in our arguments, the estimate for $\hat{\kappa}$ is in corollary 2.9. Specifically, for $\frac{1}{25} < \rho < \frac{1}{2}$, let $\kappa_1 := \hat{\kappa} - \rho$. We interpret κ_1 as a root of the equation $d(\kappa) = c_\rho^2$ for some wave speed c_ρ. Denote the two other roots of this equation κ_0 and κ_2, such that $\kappa_0 < \kappa_1 < \kappa_2$. See Figure 2-2.

The nonexistence result of this chapter can be stated as follows.

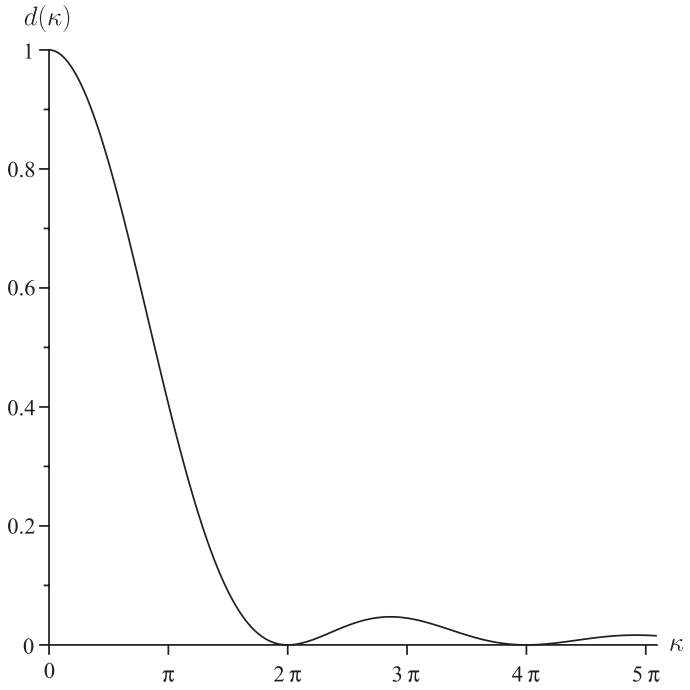

Figure 2-1: *Graph of* $d(\kappa)$ *for* $0 \leq \kappa \leq 5\pi$

Theorem 2.1. *For wave speeds* c_ρ^2 *with* $\frac{1}{25} < \rho < \frac{1}{2}$*, and* V *as in* (2.4)*, there is no travelling wave solving* (2.3) *that satisfies the single transition property* (SC) *and has bounded strain.*

One can estimate numerically that the values of c^2 for which Theorem 2.1 holds is $[0.04420, 0.04717]$, a subset of the range where the nonexistence result holds, this is shown in lemma 2.11, section 2.7. Before giving an outline proof we make the following observation. If a function $\varepsilon\colon \mathbb{R} \to \mathbb{R}$ satisfies (SC) then

$$f(x) := \Delta_1 H(\varepsilon) = \begin{cases} 1 & \text{for } x \in (-1, 0), \\ -1 & \text{for } x \in (0, 1), \\ 0 & \text{else.} \end{cases} \tag{2.12}$$

Consequently by assuming the sign condition we may reduce the nonlinear right-hand side of (2.8) into a function depending just on x and so any solution of (2.8) satisfying the sign condition (SC) also satisfies the inhomogeneous equation

$$L_c \varepsilon(x) = -2f(x). \tag{2.13}$$

We note here that since f is piecewise constant and compactly supported it has a Fourier

sine transform that can be calculated to be

$$\mathscr{F}_s[f](\kappa) = -\frac{1}{\sqrt{2\pi}} \frac{4\sin^2\left(\frac{1}{2}\kappa\right)}{\kappa}. \tag{2.14}$$

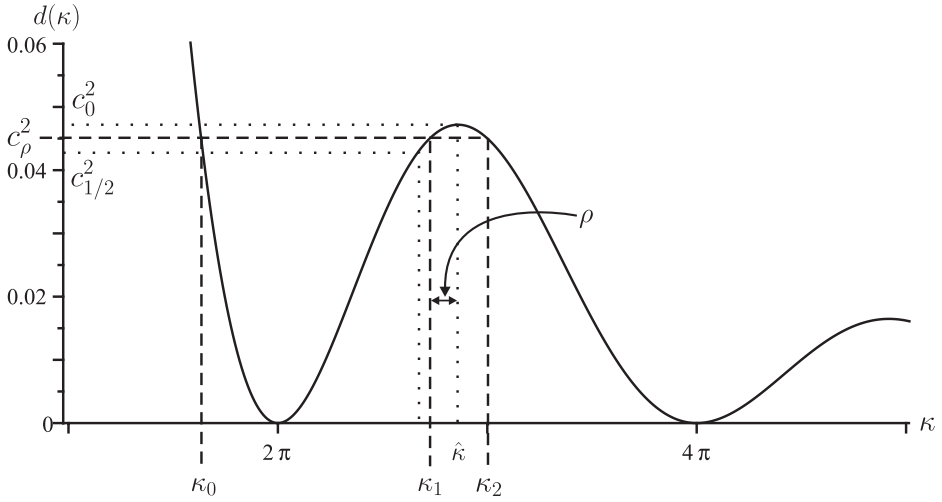

Figure 2-2: *Key notation for this chapter*

The proof outline is as follows. Fix $\frac{1}{25} < \rho < \frac{1}{2}$, this in turn determines c^2. Assume for contradiction that there exists a solution (2.1) that satisfies (SC), for the given c^2. The first step is to show that equation (2.13) has a solution. Secondly we then need to demonstrate that the solution we find violates (SC) and therefore cannot be a solution of the full equation. In a final step, since the solution we find in the first step is not unique, we demonstrate that any other distributional solution to (2.13) also fails (SC).

2.4 Profile-Corrector Method

The profile-corrector method in [SZ09b] works as follows. Define an explicit *profile function*, called $\varepsilon_{\mathrm{pr}}$, that is designed to remove the singularities in

$$\frac{\mathscr{F}[f](\kappa)}{D(\kappa, c)}. \tag{2.15}$$

Then show that $\varepsilon_{\mathrm{pr}}$ satisfies

$$L_c \varepsilon_{\mathrm{pr}}(x) = -2f(x) + \Phi(x), \tag{2.16}$$

where $\Phi \in L^2(\mathbb{R})$. We then define the *corrector function*, denoted by ε_{cor}, as the solution to

$$L_c \varepsilon_{\text{cor}}(x) = \Phi(x). \tag{2.17}$$

Then $\varepsilon := \varepsilon_{\text{pr}} - \varepsilon_{\text{cor}}$ obviously solves (2.13). The advantage now being that Φ has much better properties than $-2f$, in particular, its Fourier transform has the same zeros as $D(\kappa, c)$ and hence no singularities. We may then demonstrate failure of the sign condition (SC) as follows. First we identify some points of the profile function where the sign condition is violated. Then we show that the $L^\infty(\mathbb{R})$ norm of ε_{cor} is sufficiently small as to not change the sign of ε in the neighbourhood of the points found in the first step.

The problem of integrating over singularities induced by zeros of the dispersion relation is acknowledged in the physics literature. A *causality principle for steady-state* solution is introduced as a formal solution method [Sle02]. In this approach one integrates along paths in the complex plane that avoid the singularities in solving (2.15) then considers the limit as the path returns to that traversing the real line. The difficulty of this approach is that the representation of the solution as a formal sum makes verification of the sign condition difficult. Should the sign condition hold then as in [SZ09b] one would have existence, however, as shown here, failure of the sign condition implies nonexistence.

We define the profile function as follows. Suppose we have selected $\rho \in (\frac{1}{25}, \frac{1}{2})$ and obtained the wave speed c_ρ and the roots κ_i of $d(\kappa) = c_\rho^2$ for $i = 0, 1, 2$. Let α_i and $\beta_i > 0$ be real constants for $i = 0, 1, 2$ to be fixed later.

Adapting the approach of [SZ09b] we define a profile function as follows. First let us introduce an oscillating part as

$$\varepsilon_{\text{pr}}^{\text{osc}}(x) := \text{sign}(x) \left[\sum_{i=0}^{2} \alpha_i \left(\frac{2\sin^2\left(\frac{1}{2}\kappa_i x\right)}{\kappa_i^2} + \frac{1 - \exp(-\beta_i |x|)}{\beta_i^2} \right) \right]. \tag{2.18}$$

The purpose of $\varepsilon_{\text{pr}}^{\text{osc}}$ is to capture the oscillating tails of the solution and join them smoothly at the origin. Note that $\varepsilon_{\text{pr}}^{\text{osc}} \in C^2(\mathbb{R})$ for all values of κ_i and β_i, $i = 0, 1, 2$. We then define the jump part of the profile,

$$\varepsilon_{\text{pr}}^{\text{jump}}(x) := -\frac{1}{2c_\rho^2} \Delta_1 \left[\text{sign}(x) \frac{1}{4} x^2 \right]. \tag{2.19}$$

The purpose of this function is that when added to the profile it compensates the jumps that occur in the right-hand side of (2.13). We are now in a position to define the profile

function,

$$\varepsilon_{\mathrm{pr}}(x) := \varepsilon_{\mathrm{pr}}^{\mathrm{osc}}(x) + \varepsilon_{\mathrm{pr}}^{\mathrm{jump}}(x). \tag{2.20}$$

The values for α_i and β_i are determined in Lemma 2.2 a plot of $\varepsilon_{\mathrm{pr}}$ for these values is included in Figure 2-3.

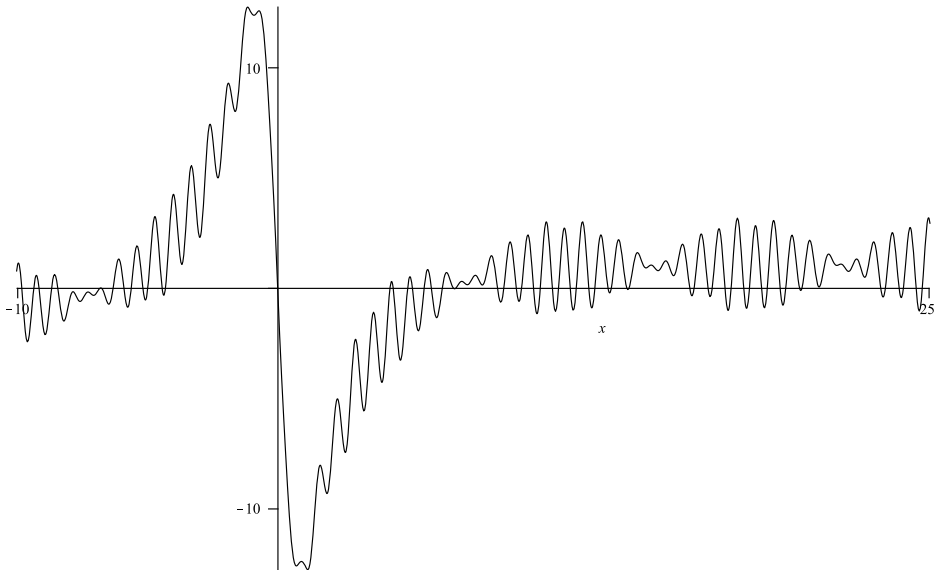

Figure 2-3: *The function ε_{pr} for $-10 < x < 25$ with wave speed $c^2 = 0.045$, illustrating the failure of* (SC)

As outlined in the introduction to this section, given the profile function defined above we need to show that there exists a function satisfying the corresponding corrector equation (2.17).

Lemma 2.2. *The profile function defined in* (2.20) *gives rise to a $\Phi \in L^2(\mathbb{R})$ as defined in* (2.16)*. Furthermore, given Φ,* (2.17) *has a unique solution in $L^2(\mathbb{R})$.*

Proof. The Fourier transform of $L_c\varepsilon_{\mathrm{pr}}$ is

$$\mathscr{F}_s[L_c\varepsilon_{\mathrm{pr}}](\kappa) = \sqrt{\frac{2}{\pi}} D(\kappa, c) \left(\sum_{i=0}^{2} \frac{\alpha_i}{\kappa(\kappa_i^2 - \kappa^2)} \frac{\beta_i^2 + \kappa_i^2}{\beta_i^2 + \kappa^2} - \frac{4\sin^2\left(\frac{1}{2}\kappa\right)}{\kappa} \frac{1}{c_\rho^2 \kappa^2} \right). \tag{2.21}$$

By (2.14), (2.16) and (2.21) it follows that

$$\begin{aligned}\mathscr{F}_s[\Phi](\kappa) &= \mathscr{F}_s[L_c\varepsilon_{\mathrm{pr}}](\kappa) + 2\mathscr{F}_s[f](\kappa)\\ &= \sqrt{\frac{2}{\pi}}\left\{D(\kappa,c)\left(\sum_{i=0}^{2}\frac{\alpha_i}{\kappa(\kappa_i^2-\kappa^2)}\frac{\beta_i^2+\kappa_i^2}{\beta_i^2+\kappa^2} - \frac{4\sin^2\left(\frac{1}{2}\kappa\right)}{c_\rho^2\kappa^3}\right) - \frac{4\sin^2\left(\frac{1}{2}\kappa\right)}{\kappa}\right\}.\end{aligned} \tag{2.22}$$

Obviously the only candidates for singularities in the Fourier transform of Φ are $\kappa \in \{0, \kappa_0, \kappa_1, \kappa_2\}$. The singularities are all removable. From these observations we conclude that the Fourier transform of Φ is bounded. Since $\mathscr{F}[\Phi] \in L^2(\mathbb{R})$ it follows from Parseval's identity that $\Phi \in L^2(\mathbb{R})$. It remains to show that, given Φ, (2.17) has a unique solution in $L^2(\mathbb{R})$. We make the following definitions

$$P(\kappa) := \prod_{j=0}^{2}(\kappa_j^2-\kappa^2) \quad \text{and} \quad p_i(\kappa) = \frac{P(\kappa)}{(\kappa_i^2-\kappa^2)} \text{ for } i = 0,1,2.$$

We also define the rescaled variables $\ell_i := \kappa/\kappa_i$. Taking the Fourier transform of (2.17) and setting

$$\gamma_i^2 := \left(1+\frac{\kappa_i^2}{\beta_i^2}\right)^{-1}$$

we find that

$$\begin{aligned}\mathscr{F}_s[\varepsilon_{\mathrm{cor}}](\kappa) &= \frac{\mathscr{F}_s[\Phi](\kappa)}{D(\kappa,c)}\\ &= \sqrt{\frac{2}{\pi}}\left\{\left(\sum_{i=0}^{2}\frac{\alpha_i}{\kappa(\kappa_i^2-\kappa^2)}\frac{\beta_i^2+\kappa_i^2}{\beta_i^2+\kappa^2} - \frac{4\sin^2\left(\frac{1}{2}\kappa\right)}{c_\rho^2\kappa^3}\right) - \frac{4\sin^2\left(\frac{1}{2}\kappa\right)}{\kappa D(\kappa,c)}\right\}\\ &= \sqrt{\frac{2}{\pi}}\left\{\frac{1}{\kappa P(\kappa)}\left(\sum_{i=0}^{2}\alpha_i p_i(\kappa)\frac{\beta_i^2+\kappa_i^2}{\beta_i^2+\kappa^2} - \frac{\left(4\sin^2\left(\frac{1}{2}\kappa\right)\right)^2}{c_\rho^2\kappa^4}\frac{\kappa^2 P(\kappa)}{D(\kappa,c)}\right)\right\}\\ &= \sqrt{\frac{2}{\pi}}\frac{1}{\kappa P(\kappa)}\left(\sum_{i=0}^{2}\frac{\alpha_i p_i(\kappa)}{\ell_i^2+\gamma_i^2(1-\ell_i^2)} - \frac{1}{c_\rho^2}\left(\frac{2\sin\left(\frac{1}{2}\kappa\right)}{\kappa}\right)^4\frac{\kappa^2 P(\kappa)}{D(\kappa,c)}\right).\end{aligned} \tag{2.23}$$

As before with the Fourier transform of Φ we see that the only candidates for singularities in (2.23) are $\kappa \in \{0, \kappa_0, \kappa_1, \kappa_2\}$. Taking the limit $\kappa \to \kappa_i$ for any $i = 0, 1, 2$ and applying l'Hôpital's rule, noting that the range of ρ ensures D has roots of single multiplicity, we

find that

$$\lim_{\kappa\to\kappa_i}\left(\sum_{i=0}^{2}\frac{\alpha_i p_i(\kappa)}{\ell_i^2+\gamma_i^2(1-\ell_i^2)}-\frac{1}{c_\rho^2}\left(\frac{2\sin\left(\frac{1}{2}\kappa\right)}{\kappa}\right)^4\frac{\kappa^2 P(\kappa)}{D(\kappa,c)}\right)= p_i(\kappa_i)\left(\alpha_i-\frac{c_\rho^2\kappa_i^3}{c_\rho^2\kappa_i-\sin(\kappa_i)}\right), \quad (2.24)$$

which vanishes if we set

$$\alpha_i := \frac{c_\rho^2\kappa_i^3}{c_\rho^2\kappa_i-\sin(\kappa_i)}. \quad (2.25)$$

The function in (2.24) therefore has a continuous extension at κ_i and in particular the continuous extension has a root at κ_i. Hence $\mathscr{F}_s[\varepsilon_{\text{cor}}]$ is bounded for $\kappa\in\{\kappa_0,\kappa_1,\kappa_2\}$. To show that $\mathscr{F}_s[\varepsilon_{\text{cor}}]$ is bounded as $\kappa\to 0$ we need to apply l'Hôpital's rule twice to find that (2.23) becomes

$$\sum_{i=0}^{2}\frac{\alpha_i}{\kappa_i^2\gamma_i^2}-\frac{1}{c_\rho^2(1-c_\rho^2)}=\sum_{i=0}^{2}\frac{\operatorname{sign}(\alpha_i)}{c_\rho^2(1-c_\rho^2)}-\frac{1}{c_\rho^2(1-c_\rho^2)}=0,$$

if we take $\beta_i>0$ to satisfy

$$\left(1+\frac{\kappa_i^2}{\beta_i^2}\right)^{-1} := |\alpha_i|\,\frac{c_\rho^2(1-c_\rho^2)}{\kappa_i^2}. \quad (2.26)$$

and the fact that $\alpha_0,\alpha_2>0$ and $\alpha_1<0$. To show that $\alpha_0,\alpha_2>0$ and $\alpha_1<0$ note that $c_\rho^2=c_\rho^2(\kappa)$, for our range of c_ρ^2 it follows that $\kappa_0\in(5.10,5.12)$, $\kappa_1\in(8.48,8.95)$ and $\kappa_2\in(9.01,9.51)$ then it is clear from a plot of α, when considered as a function of κ that this is true, a plot is included in Figure 2-4. A rigorous proof for the bounds on κ_0 and κ_2 is the subject of lemma 2.12. The bounds for κ_1 follow from corollary 2.9.

It is important to also ensure that the β_i is well defined since we also see from our plot of α that $|\alpha(\kappa)|\to\infty$ as $\kappa\to\hat{\kappa}$ therefore for sufficiently small ρ the β_i become complex. However by considering a plot, see Figure 2-5, of β as a function of κ that it is well defined for the wave speeds considered here. A rigorous proof of the fact that β_i is well defined is contained in lemma 2.14 in section 2.7.

The properties of α_i and β_i that can easily be seen by the above plots are also proved in section 2.7. We have shown that $\mathscr{F}_s[\varepsilon_{\text{cor}}]$ is bounded at all of the potential singularities and therefore bounded on $\mathbb{R}$. Since $\mathscr{F}[\varepsilon_{\text{cor}}]$ is bounded, continuous, and decays sufficiently fast at infinity, it follows that $\mathscr{F}[\varepsilon_{\text{cor}}]\in L^2(\mathbb{R})$ and hence by Parseval's identity $\varepsilon_{\text{cor}}\in L^2(\mathbb{R})$ and uniquely satisfies (2.17). □

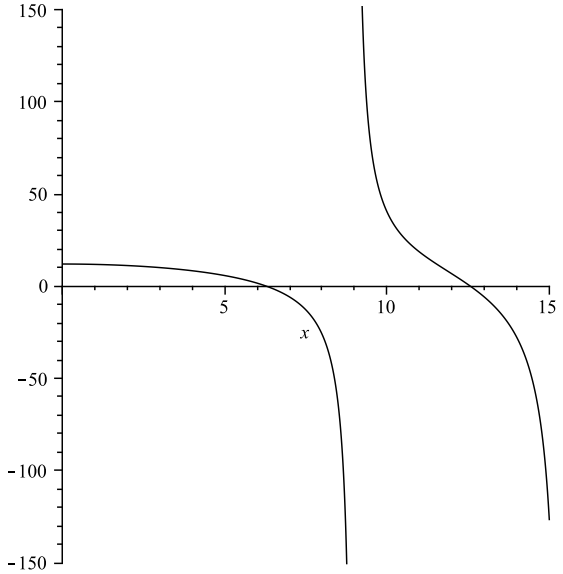

Figure 2-4: *Function α which generates the constants α_i, the asymptote is at $\hat{\kappa} \approx 8.9868$*

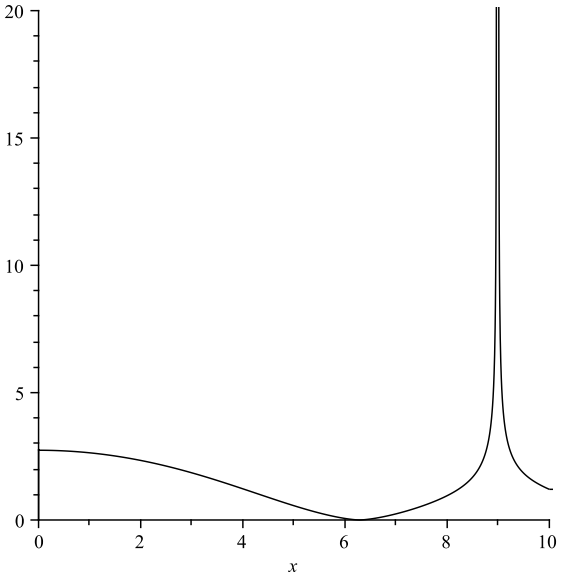

Figure 2-5: *Function β which generates the constants β_i, the discontinuity is around $\hat{\kappa} \approx 8.9868$ and lies on a set of positive measure no more than $2/25$ in diameter*

Note that in the above prove the following fact was demonstrated.

Corollary 2.3. *For every* $\frac{1}{25} < \rho < \frac{1}{2}$ *it follows that* $\alpha_0, \alpha_2 > 0$ *and* $\alpha_1 < 0$.

An alternative proof of corollary 2.3 is contained in section 2.7. The following lemma shows that the tails of the corrector decay as $x \to \pm\infty$.

Lemma 2.4. *For all* $\delta > 0$ *the set* $\{x : |\varepsilon_{cor}(x)| \geq \delta\}$ *is compact.*

Proof. We have that $\mathscr{F}[\varepsilon'_{\mathrm{cor}}](\kappa) = i\kappa\mathscr{F}[\varepsilon_{\mathrm{cor}}](\kappa) = \kappa\mathscr{F}_s[\varepsilon_{\mathrm{cor}}](\kappa)$ and therefore

$$\mathscr{F}[\varepsilon'_{\mathrm{cor}}](\kappa) = \sqrt{\frac{2}{\pi}} \left\{ \frac{1}{P(\kappa)} \left(\sum_{i=0}^{2} \frac{\alpha_i p_i(\kappa)}{\ell_i^2 + \gamma_i^2(1-\ell_i^2)} - \frac{1}{c_\rho^2} \left(\frac{2\sin\left(\frac{1}{2}\kappa\right)}{\kappa} \right)^4 \frac{\kappa^2 P(\kappa)}{D(\kappa, c)} \right) \right\}.$$

We can see that the pole at $\kappa = 0$ is removeable and the remaining potential poles are handled by the choice of α_i $i = 0, 1, 2$, as before. Then $\mathscr{F}[\varepsilon'_{\mathrm{cor}}]$ is bounded and therefore by the Plancherel theorem we have that $\varepsilon'_{\mathrm{cor}} \in L^2(\mathbb{R})$. Therefore $\varepsilon_{\mathrm{cor}} \in H^1(\mathbb{R})$ and by the Sobolev embedding theorem [Fol84, Theorem 8.54] $\varepsilon_{\mathrm{cor}}(x) \to 0$ as $x \to \pm\infty$. □

The next result determines explicitly all bounded solutions to the homogeneous version of (2.13).

Lemma 2.5. *Let* $\varepsilon \in L^\infty(\mathbb{R})$. *Then* $L_c\varepsilon = 0$ *if and only if*

$$\varepsilon \in K := \mathit{span}\left(\{1\} \cup \{\cos(\kappa_i x)\}_{i=0}^2 \cup \{\sin(\kappa_i x)\}_{i=0}^2 \right).$$

Proof. Take the Fourier transform in the context of tempered distributions. It immediately follows that since the roots of the dispersion relation are isolated that $\mathscr{F}[\varepsilon]$ is compactly supported and hence is the sum of Dirac delta. The result follows. A complete, rigorous proof is the subject of subsection 2.7.2. □

Since (2.13) is an inhomogeneous linear equation, the solution to (2.13) is only unique modulo K. From this observation it is clear that even if one shows that ε fails the sign condition (SC) then it may still satisfy it if we add a suitable combination of functions from K. Schwetlick and Zimmer show in [SZ12] that in addition to the point symmetric wave found in [SZ09b], there also exists a family of asymmetric heteroclinic travelling waves for the same range of wave speeds. This is achieved by adding suitable combinations of functions from K and showing that the sign condition (SC) is still satisfied. The next lemma demonstrates that every solution of (2.13) fails to satisfy the sign condition (SC).

The next lemma reduces the problem of determining the sign of the profile to that of studying a trigonometric polynomial for negative values. Essentially, in order to show

that the solution obtained, with any linear combination of kernel functions added violates the sign condition one needs to simply study the oscillating tail as a function extended over $\mathbb{R}$.

Note that the solution ε obtained can be expressed in the form

$$\varepsilon(x) = \operatorname{sign}(x)\varepsilon_{\text{tail}}(x) + \varepsilon_{\text{decay}}(x)$$

where $\varepsilon_{\text{decay}}(x) \to 0$ as $x \to \pm\infty$ and

$$\varepsilon_{\text{tail}}(x) := \sum_{i=0}^{2} \alpha_i \left(\frac{1}{\kappa_i^2} + \frac{1}{\beta_i^2} \right) - \sum_{i=0}^{2} \frac{\alpha_i}{\kappa_i^2} \cos(\kappa_i x) - \frac{1}{c_\rho^2}.$$

The following lemma holds.

Lemma 2.6. *Suppose that there exists a point $x \in \mathbb{R}$ where $\varepsilon_{tail}(x) < -\frac{1}{10}$. Then for any $\eta \in K$ (defined in Lemma 2.5) one of the following holds:*

(a) there exists a sequence $z_n \to \infty$ such that $\varepsilon_{tail}(z_n) + \eta(z_n) < -\frac{1}{20}$, or,

(b) there exists a sequence $z_n \to -\infty$ such that $-\varepsilon_{tail}(z_n) + \eta(z_n) > \frac{1}{20}$.

Proof. Suppose for contradiction that both (a) and (b) are not satisfied. Then there exist x, y such that $-\infty < y < 0 < x < \infty$, $\varepsilon_{\text{tail}}(z) + \eta(z) \geq -\frac{1}{20}$ for every $z > x$ and $-\varepsilon_{\text{tail}}(z) + \eta(z) \leq \frac{1}{20}$ for every $z < y$. Since $\pm\varepsilon_{\text{tail}} + \eta$ is quasi-periodic it follows that $\varepsilon_{\text{tail}}(z) + \eta(z) \geq -\frac{1}{20}$ and $-\varepsilon_{\text{tail}}(z) + \eta(z) \leq \frac{1}{20}$ for all $z \in \mathbb{R}$. Consequently $\varepsilon_{\text{tail}} \geq -\frac{1}{20}$ for all $z \in \mathbb{R}$, a contradiction to the hypotheses of the lemma. □

The proof that $\varepsilon_{\text{tail}}$ attains a negative value is contained in Section 2.5 to maintain the flow of this argument. We are now in a position to prove the main theorem.

Proof of Theorem 2.1. Fix $\frac{1}{25} < \rho < \frac{1}{2}$ and suppose the solution ε to (2.8) satisfies the sign condition (SC). Then, decomposing $\varepsilon = \varepsilon_{\text{pr}} - \varepsilon_{\text{cor}}$ with ε_{pr} as in (2.20) gives rise to a corrector function ε_{cor} by Lemma 2.2. It follows that this is only unique modulo K and find that the general solution to (2.13) is $\varepsilon + \eta$, $\eta \in K$.

By Lemma 2.4 we have that $|\varepsilon_{\text{cor}}(x)| \to 0$ as $|x| \to \infty$ so there is a $M \in \mathbb{R}$ such that if $|x| > M$ then $|\varepsilon_{\text{cor}}(x)| < \frac{1}{30}$. By Lemma 2.6 there exists a sequence $\{z_n\}_{n=1}^{\infty} \subset \mathbb{R}$ with $|z_n| \to \infty$ as $n \to \infty$ such that either $\varepsilon_{\text{pr}}(z_n) + \eta(z_n) < -\frac{1}{20}$ or $\varepsilon_{\text{pr}}(z_n) + \eta(z_n) > \frac{1}{20}$ for

each $n \in \mathbb{N}$. Choose N sufficiently large so that $|z_N| > M$. Then either

$$\varepsilon(z_N) < |\varepsilon_{\text{cor}}(z_N)| - \tfrac{1}{20} < -\tfrac{1}{60} \quad \text{if } z_N > 0$$

or

$$\varepsilon(z_N) > -|\varepsilon_{\text{cor}}(z_N)| + \tfrac{1}{20} > \tfrac{1}{60} \quad \text{if } z_N < 0.$$

Therefore for each solution of (2.13) we can find a point where the sign condition (SC) is not satisfied. This contradicts the assumption that the sign condition holds. □

2.5 Sign Failure of the Profile

The purpose of this section is to show that $\varepsilon_{\text{tail}}$ attains a negative value. Let $\frac{1}{25} < \rho < \frac{1}{2}$ and define the averages and differences of κ_i and θ_i $(i = 1, 2)$ as

$$\kappa_\sigma := \frac{\kappa_2 + \kappa_1}{2}, \quad \kappa_\delta := \frac{\kappa_2 - \kappa_1}{2}. \tag{2.27}$$

Then the following proposition holds.

Proposition 2.7. *There exists a point $X \in \mathbb{R}$ such that $\varepsilon_{tail}(X) < -\frac{1}{10}$.*

Proof. By simple manipulation,

$$\begin{aligned}\varepsilon_{\text{tail}}(x) &= \sum_{i=0}^{2} \alpha_i \left(\frac{1}{\kappa_i^2} + \frac{1}{\beta_i^2}\right) - \sum_{i=0}^{2} \frac{\alpha_i}{\kappa_i^2} \cos(\kappa_i x) - \frac{1}{c_\rho^2} \\ &= \sum_{i=0}^{2} \alpha_i \left(\frac{1}{\kappa_i^2} + \frac{1}{\beta_i^2}\right) + \sum_{i=0}^{2} \left|\frac{\alpha_i}{\kappa_i^2}\right| \cos(\kappa_i x + \theta_i) - \frac{1}{c_\rho^2}\end{aligned} \tag{2.28}$$

where $\theta_0, \theta_2 = -\pi$ and $\theta_1 = 0$ taking into account Corollary 2.3. Substituting (2.26) into (2.28), it follows that

$$\varepsilon_{\text{pr}}(x) = \sum_{i=0}^{2} \left(\frac{\alpha_i}{|\alpha_i|\, c_\rho^2(1 - c_\rho^2)}\right) + \sum_{i=0}^{2} \left|\frac{\alpha_i}{\kappa_i^2}\right| \cos(\kappa_i x + \theta_i) - \frac{1}{c_\rho^2}$$

and again by Corollary 2.3

$$\varepsilon_{\text{pr}}(x) = \frac{1}{1 - c_\rho^2} + \sum_{i=0}^{2} \left|\frac{\alpha_i}{\kappa_i^2}\right| \cos(\kappa_i x + \theta_i). \tag{2.29}$$

After some further trigonometric manipulation and using the definitions in (2.27), (2.29)

becomes

$$\begin{aligned}\varepsilon_{\mathrm{pr}}(x) = \frac{1}{1-c_\rho^2} &+ \left|\frac{\alpha_0}{\kappa_0^2}\right| \cos(\kappa_0 x - \pi) \\ &+ \left(\left|\frac{\alpha_1}{\kappa_1^2}\right| + \left|\frac{\alpha_2}{\kappa_2^2}\right|\right) \cos(\kappa_\sigma x - \tfrac{\pi}{2}) \cos(\kappa_\delta x - \tfrac{\pi}{2}) \\ &+ \left(\left|\frac{\alpha_1}{\kappa_1^2}\right| - \left|\frac{\alpha_2}{\kappa_2^2}\right|\right) \sin(\kappa_\sigma x - \tfrac{\pi}{2}) \sin(\kappa_\delta x - \tfrac{\pi}{2}). \end{aligned} \tag{2.30}$$

Suppose for now that there exists a point X where the following holds: $\left|\alpha_0/\kappa_0^2\right| \cos(\kappa_0 X - \pi) \leq 0$, $\cos(\kappa_\sigma X - \frac{\pi}{2}) = 1$, and the point X is within a distance of $4\pi/\kappa_\sigma$ of a minimum point of $\cos(\kappa_\delta x - \frac{\pi}{2})$. Evaluating (2.30) at X we find that

$$\varepsilon_{\mathrm{pr}}(X) = \frac{1}{1-c_\rho^2} + \left(\left|\frac{\alpha_1}{\kappa_1^2}\right| + \left|\frac{\alpha_2}{\kappa_2^2}\right|\right) \cos(\kappa_\delta X - \tfrac{\pi}{2}). \tag{2.31}$$

The term containing the product of sines vanishes due to the choice of X. Using a second order Taylor expansion of $\cos(\kappa_\delta x - \frac{\pi}{2})$ around the minimum point y and the fact that $|y - X| \leq 4\pi/\kappa_\sigma$, it follows that

$$\begin{aligned}\varepsilon_{\mathrm{pr}}(z) &= \frac{1}{1-c_\rho^2} - \left(\left|\frac{\alpha_1}{\kappa_1^2}\right| + \left|\frac{\alpha_2}{\kappa_2^2}\right|\right) \left(1 - \tfrac{1}{2}(y-X)^2\right) \\ &\leq \frac{1}{1-c_\rho^2} - \left(\left|\frac{\alpha_1}{\kappa_1^2}\right| + \left|\frac{\alpha_2}{\kappa_2^2}\right|\right) \left(1 - 8\pi^2 \frac{\kappa_\delta^2}{\kappa_\sigma^2}\right). \end{aligned} \tag{2.32}$$

Hence the result for X follows if

$$\frac{1}{1-c_\rho^2} - \left(\left|\frac{\alpha_1}{\kappa_1^2}\right| + \left|\frac{\alpha_2}{\kappa_2^2}\right|\right) \left(1 - 8\pi^2 \frac{\kappa_\delta^2}{\kappa_\sigma^2}\right) \leq -\frac{1}{10}.$$

Or equivalently,

$$\left(1 - 8\pi^2 \frac{\kappa_\delta^2}{\kappa_\sigma^2}\right)^{-1} \left(\frac{1}{1-c_\rho^2} + \frac{1}{10}\right) \leq \left|\frac{\alpha_1}{\kappa_1^2}\right| + \left|\frac{\alpha_2}{\kappa_2^2}\right|, \tag{2.33}$$

(note that $8\pi^2(\kappa_\delta^2/\kappa_\sigma^2) \ll 1$ by earlier considerations in Lemma 2.2). A calculation shows that

$$\left(1 - 8\pi^2 \frac{\kappa_\delta^2}{\kappa_\sigma^2}\right)^{-1} \left(\frac{1}{1-c_\rho^2} + \frac{1}{10}\right) < 1.57. \tag{2.34}$$

The numerical upper bound follows by using the numerical bounds on the roots of the dispersion relation in Lemma 2.2. Similarly it is clear from Figure 2-4 that α_i/κ_i^2 is

monotonic, and hence by the same numerical bounds it follows that

$$\left|\frac{\alpha_1}{\kappa_1^2}\right| + \left|\frac{\alpha_2}{\kappa_2^2}\right| > 1.60.$$

Alternatively, a rigorous proof of this monotonicity can be found in corollary 2.13, section 2.7. Hence (2.33) holds.

It remains to show that the point X exists. Let

$$x := \frac{\pi}{2\kappa_\delta} \text{ and } z_n := \frac{2\pi}{\kappa_\sigma} n + \frac{\pi}{2\kappa_\sigma}, \text{ for } n \in \mathbb{N}.$$

It is clear that $\cos(\kappa_\delta x - \frac{\pi}{2}) = -1$ and $\cos(\kappa_\sigma z_n - \frac{\pi}{2}) = 1$ for every n. Since $\cos(\kappa_\sigma x - \frac{\pi}{2})$ is $2\pi/\kappa_\sigma$-periodic that there exists $m \in \mathbb{N}$ such that $0 \leq x - z_m < 2\pi/\kappa_\sigma$. See Figure 2-6 for a diagrammatic explanation of the notation; the solid and dashed intervals at the bottom indicates the intervals where $\left|\alpha_0/\kappa_0^2\right| \cos(\kappa_0 x - \pi)$ has a fixed sign and the dashed curve is $\left(\left|\alpha_1/\kappa_1^2\right| + \left|\alpha_2/\kappa_2^2\right|\right) \cos\left(\kappa_\delta x - \frac{\pi}{2}\right)$. Furthermore, it is obvious that $0 \leq z_{m+1} - x < 2\pi/\kappa_\sigma$ and $2\pi/\kappa_\sigma \leq z_{m+2} - x < 4\pi/\kappa_\sigma$. It remains to show that there exists an $X \in \{z_m, z_{m+1}, z_{m+2}\}$ such that $\left|\alpha_0/\kappa_0^2\right| \cos(\kappa_0 X - \pi) \leq 0$.

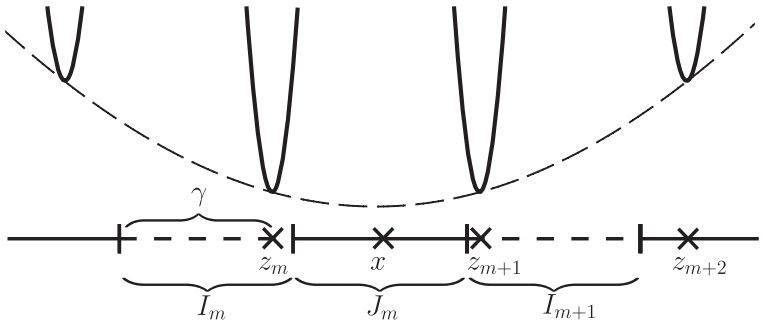

Figure 2-6: *The notation used in the proof of Proposition 2.7*

If $\left|\alpha_0/\kappa_0^2\right| \cos(\kappa_0 z_m - \pi) \leq 0$ then no further work is required. Otherwise one concludes that there exists $p \in \mathbb{N}$ such that

$$z_m = \frac{1}{\kappa_0}\left(\frac{\pi}{2} + \pi\right) + \frac{\pi(2p+1)}{\kappa_0} + \gamma,$$

for $\gamma \in (0, \pi/\kappa_0)$. This holds since we can write $(0, \infty) = (\cup_{q \in \mathbb{N}_0} I_q) \cup (\cup_{q \in \mathbb{N}_0} J_q) \cup I$, where

$$I := \left(0, \frac{3\pi}{2\kappa_0}\right), \qquad I_q := \frac{\pi(2q+1)}{\kappa_0} + \frac{3\pi}{2\kappa_0} + \left(0, \frac{\pi}{\kappa_0}\right)$$

and

$$J_q := \frac{2\pi q}{\kappa_0} + \frac{3\pi}{2\kappa_0} + \left[0, \frac{\pi}{\kappa_0}\right].$$

A simple calculation demonstrates that $\cos(x) > 0$ on I_q and $\cos(x) \leq 0$ on J_q. Since, by definition,

$$\begin{aligned} z_{m+1} &= z_m + \frac{2\pi}{\kappa_\sigma} \\ &= \frac{3\pi}{2\kappa_0} + \frac{2\pi(p+1)}{\kappa_0} + \gamma + \frac{2\pi}{\kappa_\sigma} - \frac{\pi}{\kappa_0}, \end{aligned}$$

it follows that $\left|\alpha_0/\kappa_0^2\right| \cos(\kappa_0 z_{m+1} - \pi) \leq 0$, or equivalently $z_{m+1} \in J_{p+1}$, if

$$0 \leq \gamma + \frac{2\pi}{\kappa_\sigma} - \frac{\pi}{\kappa_0} \leq \frac{\pi}{\kappa_0}. \tag{2.35}$$

Since we have explicit bounds for γ, κ_0 and κ_σ from the considerations in Lemma 2.2 a calculation shows that the lower bound in (2.35) holds uniformly in ρ. The upper bound is not necessarily satisfied and therefore we can only be sure that $\left|\alpha_0/\kappa_0^2\right| \cos(\kappa_0 z_{m+1} - \pi) \leq 0$ if $\gamma \leq 2\pi/\kappa_0 - 2\pi/\kappa_\sigma$. If we know $\gamma \leq 2\pi/\kappa_0 - 2\pi/\kappa_\sigma$ then we have found the required point, otherwise $2\pi/\kappa_0 - 2\pi/\kappa_\sigma < \gamma < \pi/\kappa_0$, the upper bound arising from the definition of γ. By definition,

$$\begin{aligned} z_{m+2} &= z_m + \frac{4\pi}{\kappa_\sigma} \\ &= \frac{2\pi(p+2)}{\kappa_0} + \gamma + \frac{4\pi}{\kappa_\sigma} - \frac{3\pi}{2\kappa_0}. \end{aligned}$$

Proceeding as before, we have that it follows that $\left|\alpha_0/\kappa_0^2\right| \cos(\kappa_0 z_{m+2} - \pi) \leq 0$, equivalently $z_{m+2} \in J_{p+2}$, if

$$0 \leq \gamma + \frac{4\pi}{\kappa_\sigma} - \frac{3\pi}{\kappa_0} \leq \frac{\pi}{\kappa_0}. \tag{2.36}$$

which is equivalent to

$$\frac{3\pi}{\kappa_0} - \frac{4\pi}{\kappa_\sigma} < \gamma < \frac{4\pi}{\kappa_0} - \frac{4\pi}{\kappa_\sigma}. \tag{2.37}$$

Using the numerical bounds on the roots of the dispersion relation from Lemma 2.2 and the assumption that $2\pi/\kappa_0 - 2\pi/\kappa_\sigma < \gamma < \pi/\kappa_0$ one can show that (2.37) holds. What we have demonstrated is that there is at least one point in $\{z_m, z_{m+1}, z_{m+2}\}$ such that $\left|\alpha_0/\kappa_0^2\right| \cos(\kappa_0 z_{m+i} - \pi) \leq 0$. Denote this point as X. □

2.6 Discussion

Here we have demonstrated that at wave speeds much less than the speed of sound, there are no travelling wave solutions that have bounded strain making a single transition between harmonic potential wells. In particular, we have shown that the solutions obtained in [SZ09b, SZ12] do not exist for the chosen significantly lower wave speeds. This confirms that for this model, the conjecture by Peyrard and Kruskal in [PK84] holds true and falls in line with the experimental observations of Förster and Scheil [FS40].

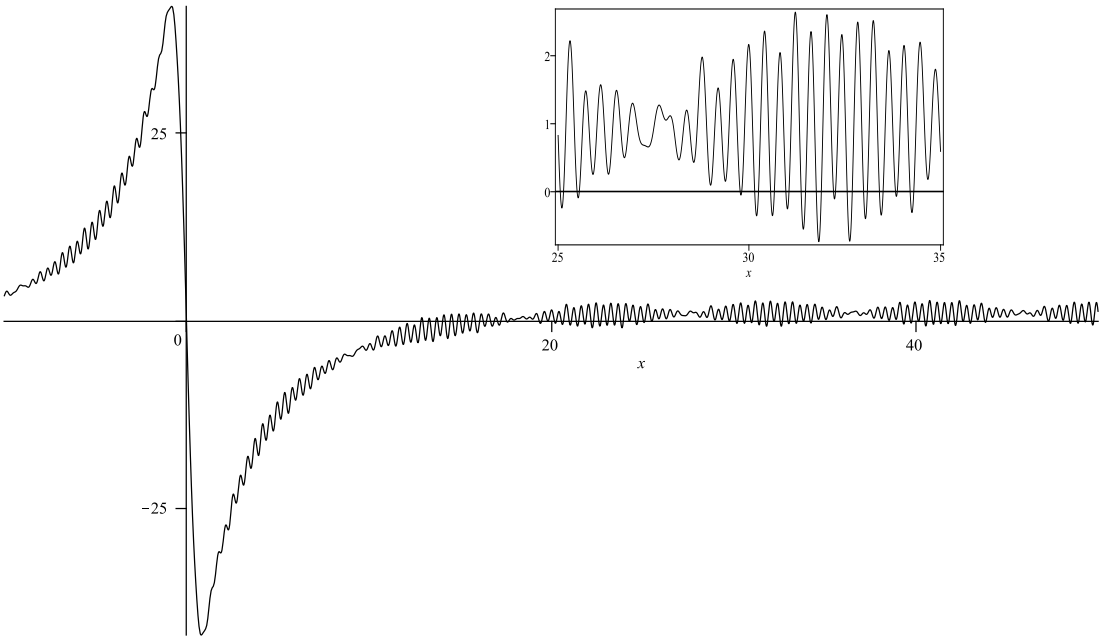

Figure 2-7: *The function ε_{pr} for $-10 < x < 50$ with wave speed $c^2 = 0.016$. (Inset) A closer view for $25 < x < 35$ illustrating the failure of* (SC)

The main feature of the proof is that as $\rho \to 0$, $|\kappa_1 - \kappa_2| \to 0$; subsequently the contributions from the kernel function resonate, causing the failure of the sign condition. One can show that Lemmata 2.2, 2.4 and 2.5 hold when $D(\kappa, c)$ has an arbitrary number of roots, with obvious modifications. The key difficulty to determining a rigorous proof for lower wave speeds is showing the equivalent of Lemma 2.6, due to the lack of information regarding the commensurability of the roots of the dispersion relation. Specifically, should one be able to prove that the set of positive roots to the dispersion relation is linearly independent over the integers then one can prove an analogue of 2.6 using Kronecker's Theorem for simultaneous Diophantine approximation [Apo90, Sections 7.4 and 7.5]. By studying the profile function numerically for wave speeds corresponding to more than three roots we observe that the nonexistence of heteroclinic travelling waves per-

sists. For instance, Figure 2-7 contains a plot of the case when $c^2 = 0.016$, a wave speed that corresponds to 5 distinct roots. The amplitude of the solution is much smaller than in the three root case, however, by close examination of the solution (see Figure 2-7 inset) one can see that the failure of the sign condition remains to hold. The failure of the sign condition extending to infinitely many points arbitrarily far away by quasi-periodicity.

It may be possible that a certain combination of kernel functions, once added to a generalised version of the corresponding profile function, cancel the resonances generated and enable the existence of a single interface travelling wave solution. We expect, however, for wave speeds close to those corresponding to a double zero of the dispersion relation that this is not the case, as we have seen here. Should one be able to prove this then one would find that there exists a sequence of intervals converging to 0 such that the same type of nonexistence result we obtain holds.

2.7 Auxiliary Proofs

In this section we prove some auxiliary results necessary for the proof of the main theorem.

2.7.1 Real Roots of the Dispersion Relation

In order to determine the failure of the sign condition (SC) it proves useful to examine the dispersion relation in greater detail. It was mentioned in Section 2.3 that by setting the parameter $\frac{1}{25} < \rho < \frac{1}{2}$ and defining $\kappa_1 := \hat{\kappa} - \rho$ we obtain three roots κ_0, κ_1 and κ_2 of the dispersion relation. This section is dedicated to proving rigorously that this is indeed the case. We will determine uniform bounds on the values which the roots could attain and estimates for the corresponding range of wave speeds.

The following lemmas give a rigorous description of what can be seen in Figures 2-1 and 2-2. In Lemma 2.8 it is demonstrated that d is monotonic on intervals that contain the roots of the dispersion relation. The existence and uniqueness of $\hat{\kappa}$ is also shown, an important result as we perform the subsequent analysis relative to this quantity. Using this one can compute bounds on important quantities such as the wave speeds for the parameter range (Lemma 2.11) and the values of the additional roots (Lemmata 2.9 and 2.12). The bounds we compute are essential for determining the sign of the profile in Proposition 2.7. Furthermore, it is demonstrated that for the parameter range we have exactly three real roots (Lemma 2.10). Once the structure of the roots is rigorously developed we are then able to determine properties of functions evaluated at these roots (Corollaries 2.3 and 2.13).

Lemma 2.8. *The function d has a unique maximum, $\hat{\kappa}$, on $(2\pi, 4\pi)$ and d is strictly monotonic decreasing on $(0, 2\pi)$ and $(\hat{\kappa}, 4\pi)$. Furthermore, d is strictly monotonic increasing on $(2\pi, \hat{\kappa})$.*

Proof. Calculating d' one finds,

$$d'(\kappa) = \frac{8}{\kappa^3}\sin^2\left(\tfrac{1}{2}\kappa\right)\left(\tfrac{1}{2}\kappa\cot\left(\tfrac{1}{2}\kappa\right) - 1\right) =: \frac{8}{\kappa^3}\sin^2\left(\tfrac{1}{2}\kappa\right)\psi(\kappa). \tag{2.38}$$

So obviously the sign of $d'(\kappa)$ equals the sign of $\psi(\kappa)$ for $\kappa > 0$. The function ψ is well defined and continuous on $\mathbb{R}^+ \setminus \{2\pi n\}_{n=0}^{\infty}$. A calculation shows that

$$\psi'(\kappa) = -\frac{1}{2}\cdot\frac{\kappa - \sin(\kappa)}{1 - \cos(\kappa)} < 0 \quad \text{for} \quad \kappa \in \mathbb{R}^+ \setminus \{2\pi n\}_{n=0}^{\infty}. \tag{2.39}$$

It follows from (2.39) that ψ is strictly monotonically decreasing on each connected component of its domain. On $(0, 2\pi)$ since ψ is continuous, strictly monotonically decreasing and $\lim_{\kappa \searrow 0}\psi(\kappa) = 0$, $\lim_{\kappa \nearrow 2\pi}\psi(\kappa) = -\infty$ it follows that ψ is negative. Similarly on $(2\pi, 4\pi)$ since ψ is continuous, strictly monotonically decreasing and $\lim_{\kappa \searrow 2\pi}\psi(\kappa) = \infty$, $\lim_{\kappa \nearrow 4\pi}\psi(\kappa) = -\infty$ it follows that there exists a unique point $\hat{\kappa}$ such that $\psi(\hat{\kappa}) = 0$. Furthermore we have that ψ is positive on $(2\pi, \hat{\kappa})$ and negative on $(\hat{\kappa}, 4\pi)$. The monotonicity of d follows as a consequence of (2.38). □

We now provide a bound for $\hat{\kappa}$ in terms of two constants $\hat{\kappa}^-$ and $\hat{\kappa}^+$.

Corollary 2.9. *It holds that $\hat{\kappa}^- < \hat{\kappa} < \hat{\kappa}^+$ where $\hat{\kappa}^{\pm} := 8.9868 \pm 10^{-4}$.*

Proof. A calculation demonstrates that $d'(\hat{\kappa}^-) > 0$ and $d'(\hat{\kappa}^+) < 0$. So by the uniqueness of $\hat{\kappa}$ on $(2\pi, 4\pi)$ by Lemma 2.8 we have $\hat{\kappa}^- < \hat{\kappa} < \hat{\kappa}^+$. □

Lemma 2.10. *For $\frac{1}{25} < \rho < \frac{1}{2}$ the equation $d(\kappa) = c_\rho^2$ has three solutions in $(0, \infty)$ that correspond to simple roots of the equation $D(\kappa) = 0$.*

Proof. Recall that for a fixed wave speed c^2 we have that κ is a root of $D(\kappa) = 0$ if and only if $d(\kappa) = c^2$. By Lemma 2.8 $d((0, 2\pi)) = (0, 1)$, $d((2\pi, \hat{\kappa})) = (0, c_0^2)$ and $d((\hat{\kappa}, 4\pi)) = (0, c_0^2)$, cf. Figures 2-1 and 2-2. Note also that d is injective when restricted to $(0, 2\pi)$, $(2\pi, \hat{\kappa})$ and $(\hat{\kappa}, 4\pi)$. Therefore given any $c^2 \in (0, c_0^2)$ it follows that the equation $d(\kappa) = c^2$ has at least three solutions $\kappa_0 \in (0, 2\pi)$, $\kappa_1 \in (2\pi, \hat{\kappa})$ and $\kappa_2 \in (\hat{\kappa}, 4\pi)$. By strict monotonicity on each of $(0, 2\pi)$, $(2\pi, \hat{\kappa})$ and $(\hat{\kappa}, 4\pi)$ it follows that the solutions obtained here correspond to simple roots of $D(\kappa) = 0$.

Recall that $c_\rho^2 := d(\hat{\kappa} - \rho)$. By Lemma 2.8 it follows that c_ρ^2 is strictly monotonically decreasing as a function in ρ. Therefore given any $\frac{1}{25} < \rho < \frac{1}{2}$ it follows that $c_\rho^2 \in (c_{1/2}^2, c_0^2)$, cf. Figure 2-2. Therefore as $(c_{1/2}^2, c_0^2) \subset (0, c_0^2)$ we have three simple roots on

$(0, 4\pi]$. We may include the points $\{2\pi, 4\pi\}$ since these just correspond to zeroes of d. It remains to demonstrate that they are the only solutions. Observe that for $\kappa \in (4\pi, \infty)$ it holds that

$$d(\kappa) < (2\pi)^{-2} < d(\hat{\kappa}^- - \tfrac{1}{2}) < c_{1/2}^2.$$

Therefore there cannot exist any solutions to $d(\kappa) = c_\rho^2$ for $\frac{1}{25} < \rho < \frac{1}{2}$ on $(4\pi, \infty)$. □

Lemma 2.11. *Let* $\frac{1}{25} < \rho < \frac{1}{2}$. *Then* $c_-^2 < c_\rho^2 < c_+^2$, *where* $c_-^2 := 0.04420$ *and* $c_+^2 := 0.04719$.

Proof. Given any $\frac{1}{25} < \rho < \frac{1}{2}$ it follows that $c_\rho^2 \in (c_{1/2}^2, c_0^2)$. By Corollary 2.9, $\hat{\kappa}^- < \hat{\kappa} < \hat{\kappa}^+$. Recall that $c_0^2 = d(\hat{\kappa})$; in order to bound c_0^2 we bound $d(\kappa)$ on $(\hat{\kappa}^-, \hat{\kappa}^+)$. As $(\hat{\kappa}^-, \hat{\kappa}^+) \subset (2\pi, 3\pi)$ it follows that $\sin\left(\frac{1}{2}\kappa\right)$ is strictly monotonic decreasing therefore $4\sin^2\left(\frac{1}{2}\kappa\right) < 4\sin^2\left(\frac{1}{2}\hat{\kappa}^+\right)$. A numerical comparison yields $d(\kappa) < 4\sin^2\left(\frac{1}{2}\hat{\kappa}^+\right)(\hat{\kappa}^-)^{-2} < c_+^2$ on $(\hat{\kappa}^-, \hat{\kappa}^+)$. In the proof of Lemma 2.10 it was seen that $d(\hat{\kappa}^- - \frac{1}{2}) < c_{1/2}^2$; calculating the lower bound one finds $c_-^2 < c_{1/2}^2$. □

Lemma 2.12. *For* $0 < \rho < \frac{1}{2}$ *we have that* $\kappa_0 \in (5.10, 5.12)$ *and* $\kappa_2 \in (9.01, 9.51)$.

Proof. Given any $0 < \rho < \frac{1}{2}$ it follows that $c_\rho^2 \in (c_{1/2}^2, c_0^2)$. Using the fact from Lemma 2.8 that d is strictly monotonic decreasing on $(0, 2\pi)$ it holds that $d(5.12) < d(\kappa) < d(5.10)$ for all $\kappa \in (5.10, 5.12)$. A trivial calculation verifies that $d(5.12) < c_-^2 - 10^{-4}$ and $d(5.10) > c_+^2 + 10^{-4}$. Therefore given any $0 < \rho < \frac{1}{2}$ it follows from Lemma 2.11 that $c_\rho^2 \in (d(5.12), d(5.10))$ and therefore the unique root κ_0 as found in Lemma 2.10 is in $(5.10, 5.12)$. Identical considerations prove the bounds for κ_2. □

With the properties of ψ and a bound on the locations of the roots of the dispersion relation one has sufficient information to prove Corollary 2.3.

Proof of Corollary 2.3. We may rearrange the form of α_i from (2.25) to obtain

$$\alpha_i = \kappa_i^2 \left(\frac{1}{1 - \frac{1}{2}\kappa_i \cot(\frac{1}{2}\kappa_i)} \right) = -\frac{\kappa_i^2}{\psi(\kappa_i)}, \tag{2.40}$$

where ψ is defined in (2.38). In the proof of Lemma 2.8 it was demonstrated that $\psi < 0$ on $(0, 2\pi)$ and $(\hat{\kappa}, 4\pi)$. Since by Lemma 2.12 $\kappa_0 \in (0, 2\pi)$ and $\kappa_2 \in (\hat{\kappa}, 4\pi)$ we have $\alpha_0, \alpha_2 > 0$. Similarly $\psi > 0$ on $(2\pi, \hat{\kappa})$ by Lemma 2.12 $\kappa_1 \in (2\pi, \hat{\kappa})$ and we have $\alpha_1 < 0$. □

Corollary 2.13. *As functions in* ρ, α_1/κ_1^2 *and* α_2/κ_2^2 *are strictly monotonic decreasing and increasing respectively.*

Proof. Rearrange (2.40) to obtain

$$\frac{\alpha_i}{\kappa_i^2} = -\frac{1}{\psi(\kappa_i)}.$$

The result then follows from the properties of ψ derived in Lemma (2.8). □

Lemma 2.14. *For $\frac{1}{25} < \rho < \frac{1}{2}$, β_i for $i = 0, 1, 2$ is well defined.*

Proof. It is clear from figure 2-5, by the regularity of β, that β_0 is well defined. The issue surrounding whether or not β_i is well posed is that we require that $\kappa^2/\beta_i^2 \geq 0$, which according to (2.26) is true if and only if

$$\frac{1}{c_\rho^2(1-c_\rho^2)} > \frac{|\alpha_i|}{\kappa^2}.$$

The remainder of the result is easy to prove using the monotonicity of α_i/κ^2 and the bounds on c_ρ^2 in corollary 2.13 and lemma 2.11 respectively. □

2.7.2 Essentially Bounded Solutions of the Linearised Equation

The purpose of this subsection is to prove Lemma 2.5. To do this we use tempered distributions. We first recall some basic properties.

Let $\mathscr{S}$ denote the space of complex valued rapidly decreasing test functions on $\mathbb{R}$, that is, functions υ which for all $m, n \in \mathbb{N}_0$ there exists $U_{m,n} \in \mathbb{R}$ such that

$$\left|\kappa^m \upsilon^{(n)}(\kappa)\right| \leq U_{m,n} \tag{2.41}$$

for all $\kappa \in \mathbb{R}$. We denote by $\mathscr{S}'$ the space of *tempered distributions*, that is, the space of linear sequentially continuous functionals acting on $\mathscr{S}$. Denote by $\langle u, \upsilon\rangle$ the action of $u \in \mathscr{S}'$ on $\upsilon \in \mathscr{S}$. Using tempered distributions one can extend the Fourier transform as a linear mapping $\mathscr{F}\colon \mathscr{S}' \to \mathscr{S}'$, defined as

$$\langle \mathscr{F}[f], \upsilon\rangle := \langle f, \mathscr{F}[\upsilon]\rangle,$$

which is bijective. A function ψ is a *multiplier* in the space $\mathscr{S}$ if it is in $C^\infty(\mathbb{R})$ and for each $n \in \mathbb{N}_0$ there exists $M_n \in \mathbb{N}_0$ such that

$$\sup_{\kappa\in\mathbb{R}} \left| \frac{\psi^{(n)}(\kappa)}{(1+|\kappa|^2)^{M_n}} \right| < \infty. \tag{2.42}$$

The space of tempered distributions is closed under multiplication by multipliers in the

space $\mathscr{S}$ [Zem65, Section 4.3]. We denote the Dirac delta distribution by δ.

The first lemma provides a decomposition of arbitrary test functions in $\mathscr{S}$.

Lemma 2.15. *Let $a_{\pm i} := \pm\kappa_{i-1}$ for $i = 1, 2, 3$ and for convenience set $a_0 := 0$. Then $\eta \in \mathscr{S}$ has the following unique representation*

$$\eta(\kappa) = \sum_{i=-3}^{3} \eta(a_i)\lambda_i(\kappa) + \eta'(0)\kappa\lambda_0(\kappa) + \chi(\kappa) \tag{2.43}$$

where $\chi(a_i) = 0$ for $i = -3, \dots, 3$ and $\chi'(0) = 0$. Furthermore $\lambda_i \in \mathscr{S}$ for $i = -3, \dots, 3$, $\lambda_i(\kappa)$ has zeroes of at least multiplicity 2 at each a_j with $i \neq j$ and $\lambda_i(a_i) = 1$ and $\lambda_i^{(m)}(a_i) = 0$ for $i = -3, \dots, 3$ and $m = 1, 2$.

Proof. The proof is, *mutatis mutandis*, the same as in [Zem65, Section 7.10, Lemma 2] but given here for the readers convenience. It follows that χ is uniquely determined by the given η, a_i and λ_i. Furthermore $\chi \in \mathscr{S}$ since η and λ_i are. The fact that $\chi(a_i) = 0$ for $i = -3, \dots, 3$ is determined by evaluating (2.43) and $\chi'(0) = 0$ follows by differentiating (2.43). □

It proves useful to know a growth estimate for all derivatives of $1/D(\kappa)$ for κ large enough.

Lemma 2.16. *For each $n \in \mathbb{N}_0$ and each $0 < \rho < \frac{1}{2}$ there exists a $P_{n,\rho} \in \mathbb{R}$ such that*

$$\left| \left(\frac{1}{D(\kappa)} \right)^{(n)} \right| < P_{n,\rho}$$

for all $\kappa \in \mathbb{R} \setminus [-Q_\rho, Q_\rho]$, where $Q_\rho := \max\{\kappa_2, 2/c^2\} + 1$.

Proof. Faà di Bruno's formula [KP02, Theorem 1.3.2] implies

$$\left| \left(\frac{1}{D(\kappa)} \right)^{(n)} \right| \leq \sum_{r=0}^{n} \frac{r!}{|D(\kappa)^{r+1}|} \sum_{j \in W(n,r)} \binom{n}{j_1, \dots, j_{n-r+1}} \prod_{q=1}^{n-r+1} \left| \frac{D^{(q)}(\kappa)}{q!} \right|^{j_q} \tag{2.44}$$

for $\kappa \in \mathbb{R} \setminus [-\kappa_2, \kappa_2]$, where

$$W(n,r) := \left\{ j \in \mathbb{N}_0^{n-r+1} : \sum_{i=1}^{n-r+1} j_i = r \text{ and } \sum_{i=1}^{n-r+1} i j_i = n \right\}.$$

Since

$$\begin{aligned}\left|\left(4\sin^2(\tfrac{1}{2}\kappa)\right)^{(n)}\right| &\leq 4\cdot\sum_{\tau=0}^{n}\binom{n}{\tau}\left|\left(\sin(\tfrac{1}{2}\kappa)\right)^{(\tau)}\left(\sin(\tfrac{1}{2}\kappa)\right)^{(n-\tau)}\right| \\ &\leq \frac{4}{2^n}\cdot\sum_{\tau=0}^{n}\binom{n}{\tau}=4 \end{aligned} \tag{2.45}$$

for all $n\in\mathbb{N}$ and clearly $D^{(n)}(\kappa)=\left(4\sin^2(\frac{1}{2}\kappa)\right)^{(n)}$ when $n>2$, it follows that $\left|D^{(n)}(\kappa)\right|\leq 4$ when $n>2$ for all $\kappa\in\mathbb{R}$. Trivially $|D''(\kappa)|\leq 2c^2+4$ on $\mathbb{R}$. Define constants

$$C_{n,j,r}:=\binom{n}{j_1,\ldots,j_{n-r+1}}\prod_{q=2}^{n-r+1}\left(\frac{\|D^{(q)}\|_\infty}{q!}\right)^{j_q};$$

it then follows that

$$\sum_{j\in W(n,r)}\binom{n}{j_1,\ldots,j_{n-r+1}}\prod_{q=1}^{n-r+1}\left|\frac{D^{(q)}(\kappa)}{q!}\right|^{j_q}\leq\sum_{j\in W(n,r)}C_{n,j,r}\left|D'(\kappa)\right|^{j_1}. \tag{2.46}$$

The definition of $W(n,r)$ implies that $0\leq j_1\leq r$, therefore the bound in (2.46) is a polynomial of at most degree r. It is easy to verify that

$$|D(\kappa)|\geq c^2\kappa^2-4 \text{ and } \left|D'(\kappa)\right|\leq 2c^2|\kappa|+4 \tag{2.47}$$

for $\kappa\in\mathbb{R}\setminus\left([-\kappa_2,\kappa_2]\cup[-2/c^2,2/c^2]\right)$. Hence combining (2.44), (2.46) and (2.47), we conclude

$$\left|\left(\frac{1}{D(\kappa)}\right)^{(n)}\right|\leq\sum_{r=0}^{n}\sum_{j\in W(n,r)}r!\,C_{n,j,r}\frac{\left(|\kappa|+\frac{2}{c^2}\right)^{j_1}}{\left(\kappa^2-\frac{2}{c^2}\right)^{r+1}} \tag{2.48}$$

for $\kappa\in\mathbb{R}\setminus\left([-\kappa_2,\kappa_2]\cup[-2/c^2,2/c^2]\right)$. Since the summands in (2.48) are rational functions whose numerator is of a lower degree then the denominator, we may uniformly bound each summand by $A_{j,r}$ on $\mathbb{R}\setminus[-Q_\rho,Q_\rho]$. The choice of Q_ρ ensures that κ is sufficiently far away from the poles of (2.48). By setting

$$P_{n,\rho}:=\sum_{p=0}^{n}\sum_{j\in W(n,r)}r!\,C_{n,j,r}A_{j,r}$$

it follows that the result holds for $\kappa\in\mathbb{R}\setminus[-Q_\rho,Q_\rho]$. □

We now prove a further technical result enabling us to prove Lemma 2.5. Note that

$D(\kappa)$ is even and therefore if κ_i is a root then so is $-\kappa_i$.

Lemma 2.17. *The function D as defined in (2.10) is a multiplier in $\mathscr{S}$. Furthermore, for $0 < \rho < \frac{1}{2}$ and $\phi \in \mathscr{S}'$ it follows that $D\phi = 0$ if and only if*

$$\phi(\kappa) \in \mathit{span}\left(\left\{\delta(\kappa), \delta'(\kappa)\right\} \cup \left\{\delta(\kappa - \kappa_i), \delta(\kappa + \kappa_i)\right\}_{i=0}^{2}\right).$$

Proof. Obviously $D \in C^\infty(\mathbb{R})$. Observe that (2.42) holds for D and $n \geq 2$ by setting $M_n := 0$ as was demonstrated in (2.45). It is clear that (2.42) holds for $n < 2$ when setting $M_n := 1$. Therefore the dispersion relation D is in $\mathscr{S}$ and the product $D\phi$ is well defined for all $\phi \in \mathscr{S}'$.

The next step is to demonstrate that for $\chi \in \mathscr{S}$ we can write $\chi(\kappa) = D(\kappa)\upsilon(\kappa)$ for all $\kappa \in \mathbb{R}$ with $\upsilon \in \mathscr{S}$ if and only if $\chi(a_i) = 0$ for $i = -3, \ldots, 3$ and $\chi'(0) = 0$, where the a_i are defined in Lemma 2.15. Necessity in this case is clear. Conversely, suppose that $\chi(a_i) = 0$ for $\chi(a_i) = 0$ for $i = -3, \ldots, 3$ and $\chi'(0) = 0$ for some $\chi \in \mathscr{S}$. Set $\upsilon(\kappa) := \chi(\kappa)/D(\kappa)$ and note that υ is smooth away from the zeros of D. Applying Taylor's theorem with integral remainder around the 0 one has

$$\frac{\chi(\kappa)}{D(\kappa)} = \frac{1}{D(\kappa)} \int_0^\kappa (\kappa - t)\chi''(t)\,\mathrm{d}t \tag{2.49}$$

in a punctured neighbourhood U_0 of 0 which contains no other root of D. The change of variable $t \mapsto \kappa s$ simplifies (2.49) to

$$\begin{aligned}\frac{1}{D(\kappa)} \int_0^\kappa (\kappa - t)\chi''(t)\,\mathrm{d}t &= \frac{\kappa^2}{D(\kappa)} \int_0^1 (1 - s)\chi''(\kappa s)\,\mathrm{d}s \\ &= \frac{1}{f(\kappa)} \int_0^1 (1 - s)\chi''(\kappa s)\,\mathrm{d}s. \end{aligned} \tag{2.50}$$

where f is smooth and nonzero in a neighbourhood V_0 of 0. It is then clear that υ on $U_0 \cap V_0$ has a continuous extension to 0 for all of its derivatives since we can differentiate under the integral in (2.50). A similar argument shows that υ has a smooth extension at all the zeros of D, therefore $\upsilon \in C^\infty(\mathbb{R})$. To determine the decay of υ, fix $m, n \in \mathbb{N}_0$. Then

$$\begin{aligned}\left|\kappa^m \upsilon^{(n)}(\kappa)\right| &\leq \sum_{r=0}^{n} \binom{n}{r} \left|\kappa^m \chi^{(r)}(\kappa) \left(D(\kappa)^{-1}\right)^{(n-r)}\right| \\ &\leq \sum_{r=0}^{n} \binom{n}{r} C_{m,r} P_{n-r,\rho} =: T_{m,n}\end{aligned}$$

on $\mathbb{R} \setminus [-Q_\rho, Q_\rho]$ due to $\chi \in \mathscr{S}$ and Lemma 2.16. By continuity there exists $U_{m,n}$ such

that (2.41) holds on $[-Q_\rho, Q_\rho]$. Therefore each derivative of υ vanishes faster than any power of κ and hence $\upsilon \in \mathscr{S}$.

We are now able to determine the conclusion of the lemma. It is clear that sufficiency holds. Conversely, let $\upsilon \in \mathscr{S}$ be arbitrary. For the dispersion relation $D \in \mathscr{S}$ it holds that

$$\langle D\phi, \upsilon\rangle = \langle \phi, D\upsilon\rangle = 0. \tag{2.51}$$

Now applying Lemma 2.15 to any $\eta \in \mathscr{S}$ we may write

$$\langle \phi, \eta\rangle = \sum_{i=-3}^{3} \eta(a_i)\langle \phi, \lambda_i(\kappa)\rangle + \eta'(0)\langle \phi, \kappa\lambda_0(\kappa)\rangle + \langle \phi, \chi\rangle$$

for λ_i, a_i and χ as described in Lemma 2.15. Now since $\chi(a_i) = 0$ for $i = -3, \dots, 3$ and $\chi'(0) = 0$ we use the first part of the proof to write $\chi = D\upsilon$ for $\upsilon \in \mathscr{S}$, then by (2.51) it follows that $\langle \phi, \chi\rangle = 0$. Now defining, $\tilde{a}_i := \langle \phi, \lambda_i\rangle$ for $i = -3, \dots, 3$ and $\tilde{b} = \langle \phi, \kappa\lambda_i\rangle$ we see that

$$\langle \phi, \eta\rangle = \sum_{i=-3}^{3} \tilde{a}_i\eta(a_i) + \tilde{b}\eta'(0)$$

which implies $\phi(\kappa) \in \operatorname{span}\left(\{\delta(\kappa), \delta'(\kappa)\} \cup \{\delta(\kappa - \kappa_i), \delta(\kappa + \kappa_i)\}_{i=0}^{2}\right)$. □

There is now enough information to complete the proof of Lemma 2.5.

Proof of Lemma 2.5. Taking the Fourier transform of the equation for $\varepsilon \in \mathscr{S}'$ we obtain

$$D(\kappa)\mathscr{F}[\varepsilon] = 0.$$

Since $\varepsilon \in \mathscr{S}'$ it follows that $\mathscr{F}[\varepsilon] \in \mathscr{S}'$. It follows by Lemma 2.17 that ε solves $L_c\varepsilon = 0$ in $\mathscr{S}'$ if and only if

$$\mathscr{F}[\varepsilon](\kappa) = \sum_{i=-3}^{3} \tilde{a}_i\delta(\kappa - a_i) + \tilde{b}\delta'(\kappa). \tag{2.52}$$

It is clear by inversion of the Fourier transform that (2.52) holds if and only if $\varepsilon \in K \cup \operatorname{span}\{x\}$. Consequently ε solves $L_c\varepsilon = 0$ in $L^\infty(\mathbb{R})$ if and only if $\varepsilon \in K$. □

Chapter 3

Homogenisation of Metric Functionals and Hamiltonian Dynamics

3.1 Introduction

The purpose of this chapter is to study the convergence of minimisers of functionals of the form

$$F_\varepsilon(u) := \int_0^1 f_\varepsilon\left(\frac{u(\tau)}{\varepsilon}, u'(\tau)\right) \mathrm{d}\tau,\ u \in W^{1,\infty}(0,1),$$

with f_ε a Carathéodory function that is periodic in the first variable and convex and $1-$homogeneous in the second. We take $W^{1,\infty}(0,1)$ to be the space of Lipschitz curves $u\colon (0,1) \to \mathbb{R}^d$. Classically, such functionals are to be interpreted as the length of the curve u in a Finsler manifold whose geometry is characterised by the function f_ε. The function f_ε may be considered a length density in the manifold. A typical form for the function f_ε is that $f_\varepsilon(s,\xi) = a_\varepsilon(s)\|\xi\|$, in which case we say that f_ε is the length density of a Riemannian manifold with metric coefficient a_ε. For the functional F_ε to be well defined and have minimisers we will assume for each ε that there exists $\beta_\varepsilon \geq \alpha > 0$ such that

$$\alpha\|\xi\| \leq f_\varepsilon(s,\xi) \leq \beta_\varepsilon\|\xi\|.$$

There are several applications to which the study of F_ε is relevant. The first, and most prominent, is that of Fermat's principle. Fermat's principle, in its original form, states that the path taken for a ray of light to travel between two given points should minimise the total travel time. To describe Fermat's principle mathematically, one considers f_ε to be the length density of a Riemannian manifold. The refractive index of

an optical material is a relative measure of how the velocity of light changes as it penetrates the material. The function a_ε in this context specifies the refractive index of the material pointwise throughout the body. In particular we take a_ε to be defined on $\mathbb{R}^d$ and therefore our optical medium would be infinite and periodic. Given $\xi_1, \xi_2 \in \mathbb{R}^d$, minimising F_ε over the space of curves joining ξ_1 to ξ_2 is then the mathematical statement of Fermat's principle. Should such a curve exist then it is called a *geodesic*. The function F_ε is said to describe the optical path length. We remark that in this setting we only consider refracted rays of light, as a reflected ray fails to minimise the optical path length [ACM09]. Consequently, this model is at most applicable when reflected light can be ruled out. Returning to the definition of F_ε, we see that as $\varepsilon \to 0$ the optical material potentially becomes very complex over small length scales. The study and numerical computation of rays of light joining two given points in the material is a difficult and computationally challenging task. At this point it is natural to ask whether or not there is a better alternative. The first question is, for a sequence $\{\varepsilon_k\}_{k=1}^\infty$ converging to zero, whether the minimising curves of F_{ε_k} joining ξ_1 to ξ_2 converge in a reasonable sense as $k \to \infty$? Since with this knowledge we may study the limit function with the additional knowledge of how the true solutions approximate it for ε sufficiently small. Such compactness is easy to conclude should the values of the functionals F_ε converge along the minimising curves. To achieve this we can equip the functionals F_ε with a notion of convergence, called Γ-convergence, that ensures that the limit functional has the properties we seek.

3.1.1 Γ-convergence

Γ-convergence was first introduced in [DF75] as a mode of convergence for functionals where a convergent relationship between minimising points is required. Let us define Γ-convergence for functionals that depend on a continuous parameter, such as F_ε that we consider in this chapter.

Definition 3.1 (Γ-convergence)**.** *Let (X, d) be a metric space and consider a sequence of functionals $g_\varepsilon \colon X \to [0, \infty)$ for $\varepsilon > 0$. We say that g_ε $\Gamma(d)$-converges to $g \colon X \to [0, \infty)$ as $\varepsilon \to 0$ if for all $x \in X$ and every sequence $(\varepsilon_k)_{k=1}^\infty \subset (0, \infty)$ converging to 0 we have*

$$\liminf_{\varepsilon \to 0} g_{\varepsilon_k}(x_{\varepsilon_k}) \geq g(x) \text{ for every sequence } (x_{\varepsilon_k})_{k=1}^\infty \text{ converging to } x, \tag{3.1}$$

and

$$\text{there exists a sequence } (x_{\varepsilon_k})_{k=1}^\infty, \text{ converging to } x, \text{ such that } \lim_{\varepsilon \to 0} g_{\varepsilon_k}(x_{\varepsilon_k}) = g(x). \tag{3.2}$$

Condition (3.1) is typically referred to as the *lim-inf inequality*, and condition (3.2) the *lim-sup inequality*, for obvious reasons. In definition 3.1 g is called the $\Gamma(d)$-limit and we write $g(x) = \Gamma(d) - \lim_{\varepsilon\to 0} g_\varepsilon(x)$, if the topology is clear from the context then the dependence on the metric is dropped.

We also say that a sequence of functionals $(g_\varepsilon)_{\varepsilon>0}$ are *equi-mildly coercive* on X if there exists a d-compact set $K \subset X$ such that $\inf_X g_\varepsilon = \inf_K g_\varepsilon$ for all $\varepsilon > 0$. The functionals F_ε are *equi-mildly coercive*, provided the infimum is in fact a minimum. The following theorem states that should the sequence F_ε Γ-converge then limit functional encodes information about the limiting minimising curves.

Theorem 3.2 (Fundamental theorem of Γ-convergence)**.** *Let X be a topological space with topology Σ and let $F_\varepsilon\colon X \to [0,\infty)$ for $\varepsilon > 0$ be a sequence of equi-mildy coercive functions on X with Γ-limit $F\colon X \to [0,\infty)$. Then, the limit $\lim_{\varepsilon\to 0} \inf_X F_\varepsilon$ exists and is equal to $\min_X F$. Furthermore, if there exists a converging sequence $(x_{\varepsilon_k})_{k=1}^\infty$ with $\lim_{k\to\infty} F_{\varepsilon_k}(x_{\varepsilon_k}) = \lim_{k\to\infty} \inf_X F_{\varepsilon_k}$, then its limit is a minimum point of F.*

Proof. Cf. [BD98, Theorem 7.2]. □

The procedure of replacing a problem like minimising F_ε with the minimisation of a simpler problem is known as *homogenisation*. There are several books in the literature outlining several other applications and theorems for Γ-convergence; we mention [BD98, Bra02, Dal93].

We mention other techniques that have been successfully applied to other homogenisation problems, notably G-convergence [Spa68, DS73], compensated compactness [Mur78, Tar79], the perturbed test function method [Eva89], H-measures [Tar90] and two-scale convergence [All92].

3.1.2 A Metric Formulation of Hamiltonian Dynamics

Before reviewing the results in the literature concerning the homogenisation of F_ε, we review further applications. One example is to determine an effective description for the motion of a Hamiltonian particle forced by a rapidly oscillating potential. Without loss of generality, we consider a particle of unit mass. We model the rapidly oscillating forcing by a potential $V(q/\varepsilon)$, where we take $\varepsilon > 0$ small and $V\colon \mathbb{R}^N \to \mathbb{R}$ is assumed to be bounded and periodic. The equation of motion for the particle is classically determined by Newton's second law

$$\ddot{q}_\varepsilon(t) = -\frac{1}{\varepsilon}\nabla V\left(\frac{q_\varepsilon(t)}{\varepsilon}\right). \tag{3.3}$$

Here the function $q_\varepsilon\colon (0,T) \to \mathbb{R}^d$ maps time to a position in space. As discussed for Fermat's principle, we would like to determine the effective behaviour of solutions to (3.3)

as $\varepsilon \to 0$. That is, we would like to determine the limits of the sequence $\{q_\varepsilon\}_{\varepsilon>0}$, provided they exist, as an approximation of q_ε for ε small. Another reformulation of (3.3), via the Routh reduction procedure [MR99, Section 8.9] to the Jacobi metric. The *Maupertuis principle* states that solutions of (3.3) with fixed total energy E are re-parameterised critical points of the *Jacobi length functional*

$$J_\varepsilon(u) := \int_0^1 \sqrt{2\left(E - V\left(u(\tau)/\varepsilon\right)\right)}\left|u'(\tau)\right| \mathrm{d}\tau. \tag{3.4}$$

Since V is bounded we take throughout this chapter $E > \|V\|_\infty$ to ensure that J_ε is the length functional of a Riemannian manifold. To recover a solution of (3.3) from a critical point of (3.4) one simply reparameterises to physical time via the equation

$$t(\tau) := \int_0^\tau \frac{|u'(\sigma)|}{\sqrt{2(E - V(u(\sigma)/\varepsilon))}} \mathrm{d}\sigma. \tag{3.5}$$

To prove the Maupertuis principle, one shows that the Euler-Lagrange equation of (3.4) is (3.3) after using the change of variables (3.5). Applying the results concerning F_ε to (3.4) could then provide insight into the limiting trajectories of (3.3) that additionally minimise (3.4). We remark that whilst the equivalence of (3.3) and (3.4) is clear for $\varepsilon > 0$, how to recover meaningful dynamical information from the limiting procedure is not, we approach this topic in section 3.5. To the best of the authors knowledge, the Maupertuis principle has not been used to study the limiting dynamics of (3.3).

3.1.3 The Hamilton-Jacobi PDE

Finally we will discuss how the study of F_ε relates to the Hamilton-Jacobi PDE. The ε-dependant Hamilton-Jacobi PDE that we will consider takes the form

$$\partial_t w_\varepsilon(x,t) + H\left(x/\varepsilon, \nabla_x w_\varepsilon(x,t)\right) = 0 \tag{3.6}$$

subject to a suitable initial condition $w_\varepsilon(x,0) = w_0(x) \in BUC(\mathbb{R}^d)$. In this thesis we only consider systems where energy is conserved, as a consequence we only consider Hamilton-Jacobi equations where the Hamiltonian is independent of time (for time dependant Hamiltonians see [Bra92, BD98, GE02]). The function H is assumed to be convex in the second variable and is additionally superlinear, that is

$$\lim_{\|p\|\to\infty} \frac{H(x/\varepsilon, p)}{\|p\|} = \infty,$$

this is to ensure that the Legendre transform of H, defined as

$$H^*(x/\varepsilon, q) := \sup_{p\in\mathbb{R}^d} \{q \cdot p - H(x/\varepsilon, p)\}, \tag{3.7}$$

has the property that $H^{**} = H$. The function H^* is called the Lagrangian and is denoted as L. In order to speak of a homogenisation problem we suppose that H is $[0,1]^d$-periodic in the first variable, and seek to study the convergence of solutions of (3.6). In what follows we sketch a method of homogenising (3.6), for a rigourous proof using Γ-convergence see [LPV88, Bra92, BD98, E91]. The unique solution of (3.6) is given by

$$w_\varepsilon(x,t) = \inf_{y\in\mathbb{R}^d} \{w_0(y) + S_\varepsilon(x,t;y,s)\colon 0 \le s \le t\}$$

where

$$S_\varepsilon(x,t;y,s) := \inf_{q\in W^{1,\infty}(0,1)} \left\{ \int_s^t L(q(t)/\varepsilon, q'(t))\,\mathrm{d}t\colon q(s) = y, q(t) = x \right\}$$

is the Hopf-Lax formula [BD98, Eva98]. Applying Γ-convergence techniques it is possible to show that the functions S_ε converge locally uniformly to a function of the form

$$S(x,t;y,s) := \inf_{q\in W^{1,\infty}(0,1)} \left\{ \int_s^t \bar{L}(q'(t))\,\mathrm{d}t\colon q(s) = y, q(t) = x \right\}.$$

Note that this also ensures the local uniform convergence of w_ε. It is possible to show that $\bar{L}$ is superlinear and convex and therefore $\bar{L}^*$ exists and we denote it by $\bar{H}$, the *effective Hamiltonian*. Consequently S solves the Hamilton-Jacobi PDE

$$\partial_t w_\varepsilon(x,t) + \bar{H}\left(\nabla_x w_\varepsilon(x,t)\right) = 0 \tag{3.8}$$

subject to a suitable initial condition $w_\varepsilon(x,0) = w_0(x) \in BUC(\mathbb{R}^d)$. Note that equation (3.3) has the Hamiltonian

$$H(p,x) = \frac{1}{2}\|p\|^2 + V(x).$$

Consequently one could ask how the above homogenisation procedure encodes dynamical information into the effective Hamiltonian; this question is addressed in [GE01, GE02], their focus is primarily on *absolute minimisers* rather than boundary value problems that we consider here. An absolute minimising trajectory $r\colon (0,T) \to \mathbb{R}^d$ has the property that for all $S \in (0,T)$

$$\int_0^S L(r(t), r'(t))\,\mathrm{d}t \le \int_0^S L(q(t), q'(t))\,\mathrm{d}t$$

for all q such that $r(0) = q(0), r(S) = q(S)$, and is therefore not as general as considering arbitrary boundary conditions. The homogenisation of the Hamilton-Jacobi PDE may also be done using viscosity solution techniques [LPV88]. It was in the unpublished preprint [LPV88] that the homogenisation of the Hamilton-Jacobi PDE was first completed.

It would be natural to ask, how do the Hamilton-Jacobi PDE and Maupertuis principle relate in their effective description of Hamiltonian trajectories? We show in section 3.5 that in one space dimension that both approaches are equivalent, under suitable assumptions, and prove that in higher dimensions that they yield the same sets of information for boundary value problems. The connection between F_ε and the Hamilton-Jacobi PDE is a little more subtle and we return to this after reviewing the results in the literature concerning the homogenisation of F_ε.

3.1.4 Results on Metric Homogenisation

The homogenisation of metric functionals has had a relatively short history in mathematics, and to the best of the author's knowledge the main results are covered below. Initially, mathematicians were interested in the homogenisation of functionals of the form

$$\int_0^1 f\left(\frac{u(\tau)}{\varepsilon}, u'(\tau)\right) \mathrm{d}\tau, \quad u \in W^{1,p}(0,1), \tag{3.9}$$

where f is periodic in the first variable, and satisfy growth conditions of the form

$$\alpha\|\xi\| \leq f(s,\xi) \leq \beta\|\xi\|.$$

The Γ-convergence of such functionals was described by [BD78], and then developed further by [E91], in both cases for $p > 1$. The challenge that mathematicians faced when first tackling homogenisation problems of this type was the lack of periodicity for the minimisers, this meant that the cell problem formula that would arise, for instance in the homogenisation of uniformly elliptic operators, no longer existed. The main result states that the sequence of functionals Γ-converge, with respect to the norm topology on $L^p(0,1)$, to a functional of the form

$$\int_0^1 \tilde{f}(u'(\tau))\,\mathrm{d}\tau,$$

where $\tilde{f}$ would be described by an *asymptotic homogenisation formula*

$$\tilde{f}(\xi) := \lim_{T\to\infty} \inf_{w\in W_0^{1,p}(0,T)} \frac{1}{T}\int_0^T f\left(\xi\tau + w(\tau), \xi + w'(\tau)\right) \mathrm{d}\tau,$$

which essentially generalises the cell-problem formula. To see the reduction of the asymptotic homogenisation formula to a cell problem formula for convex spatially periodic problems consult [BD98, Theorem 14.7]. The extension of these results to f almost periodic followed in [Bra86] and [Bra92].

In [BD78] it is shown that the limit of Riemannian length functionals does not yield a Riemannian length functional in the limit. For a detailed example of this phenomenon in the literature see [BD98, Example 16.2], and section 3.4 in this thesis for a new example. We give an overview of [BD98, Example 16.2], there the authors had a Riemannian length density with metric coefficient

$$a(s) := \begin{cases} \alpha & \text{if } s \in \Omega := (0,1/2)\times(1/2,1) \cup (1/2,1)\times(0,1/2) \\ \beta & \text{if } s \in [0,1]^d \setminus \Omega, \end{cases} \tag{3.10}$$

and select $4\beta < \alpha$ to ensure that geodesics with endpoints in $[0,1]^d \setminus \Omega$ stay in $[0,1]^d \setminus \Omega$ (see lemma 3.21). In this case the limit length density $\tilde{f}$ is given by

$$\tilde{f}(\xi) = \beta\left((\sqrt{2}-1)\min\{\|\xi_1\|, \|\xi_2\|\} + \max\{\|\xi_1\|, \|\xi_2\|\}\right)^2, \tag{3.11}$$

the level sets of $\tilde{f}$ are regular octagons. The underlying microscopic features of the checkerboard metric make it easy to compute geodesics by elementary geometric reasoning.

The case when $p = 1$ was studied in [AV98]. In [AV98] the authors considered homogenisation on $W^{1,1}(0,1)$ with respect to the $L^1(0,1)$ topology. In addition, the authors also showed in the cases when f is 1-homogeneous in the second variable, that if (3.9) Γ-converges for $p = 1$, with respect to the $L^1(0,1)$ topology, then the functionals

$$\int_0^1 f\left(\frac{u(\tau)}{\varepsilon}, u'(\tau)\right)^r \mathrm{d}\tau, \quad u \in W^{1,r}(0,1), \tag{3.12}$$

Γ-converge, with respect to the $L^r(0,1)$ topology, to

$$\int_0^1 \tilde{f}(u'(\tau))^r \,\mathrm{d}\tau.$$

The next major step in the homogenisation of length functionals was the paper of [BPF01]

which studied the topological equivalence of problems involving distances. In particular, the authors showed that if the induced metrics

$$d_\varepsilon(\xi_1,\xi_2) := \inf_{w\in W^{1,p}(0,1)} \left\{ \int_0^1 f\left(\frac{u(\tau)}{\varepsilon}, u'(\tau)\right) \mathrm{d}\tau \colon u(0)=\xi_1, u(1)=\xi_2 \right\},$$

with $p=1$, converge locally uniformly to a metric, then the length functionals Γ-converge on $W^{1,\infty}(0,1)$ with respect to the $L^\infty(0,1)$ topology to the length functional associated to d, and vice versa. Details of this result are in section 3.2. With this theorem it was possible to conclude that the Finsler length functionals were closed under Γ-convergence and since the Riemannian metrics were not closed under Γ-convergence the authors were lead to ask whether the Riemannian metrics were dense in the class of Finsler metrics. This question was subsequently answered in [BBF02], and the conjecture shown to be true, in fact the authors showed that the class of smooth isotropic Riemannian metrics were dense in the class of Finsler metrics.

Another notable contribution to the theory of metric homogenisation is that of [DP07] where the authors considered 'two-phase' metrics. For this setting the function f is assumed to be a Riemannian length density with metric coefficient

$$a(s) := \begin{cases} \alpha & \text{if } s \in \Omega \subset [0,1]^d \\ \beta & \text{if } s \in [0,1]^d \setminus \Omega, \end{cases}$$

and extended periodically. They show that such functionals are dense in the class of Finsler metrics.

The most recent contribution, to the best of the author's knowledge, is the generalisation of [BD98, Example 16.2] in [ACM09]. There the authors considered a Riemannian length density with metric coefficient (3.10), however the value of α is to be taken lower. This complicates the computation of geodesics, therefore the authors use an approach based on Snell's law to solve the problem. The limit length density in this case is a hexadecagon [ACM09, Theorem 6.3].

3.1.5 Metric-Type Hamiltonians

Returning to the result of [AV98], it is clear that understanding the effective behaviour of the functionals (3.9) is equivalent to understanding the effective behaviour of the functionals (3.12) for $r > 1$. The latter functionals enjoy convexity and superlinear growth in the second variable, therefore they can be thought of as a Lagrangian. Hence, for the results that we obtain in this thesis we also obtain results for the homogenisation

of Hamilton-Jacobi PDE of the form

$$\partial_t w_\varepsilon(x,t) + g\left(x/\varepsilon, \nabla_x w_\varepsilon(x,t)\right) = 0, \tag{3.13}$$

where $g(x,q) = (f(x,q)^r)^*$. In particular, the example of section 3.4 has immediate implications for those attempting to study such homogenisation problems numerically as in [CB03, GO04, OTV09].

3.1.6 Results in this Chapter

This chapter of the thesis is roughly organised to make a key point in each section. In section 3.2 we revisit the classical problem of homogenising uniformly bounded metric functionals. While the homogenisation of such functionals has been studied before we combine existing results together to determine more information about the limit problem. One particular feature of this work is that the induced metrics converge locally uniformly to a norm, rather than a metric as given by the result of [BPF01].

In section 3.3 we approach the problem of homogenising metric functionals that are not uniformly bounded. Such problems have been considered before but only as particular examples for instance in [BPF01]. We restrict to the case of two-phase Riemannian length densities. As a result we show that the length functionals still Γ-converge. However, the homogenisation of the boundary value problem depends on the growth of the metric. We find that for sufficiently slow growth the boundary value problem Γ-converges, otherwise the problem fails to Γ-converge. This behaviour has been previously unobserved in the literature and provides additional insight into the result of [BPF01].

In section 3.4 we perform an explicit computation of a metric using the theory of section 3.2. The example depends on a parameter that controls the microscopic features of the ε-dependant functional, a feature that apparently has not previously existed in the literature. Furthermore, this is apparently the first documented instance of a limit metric consisting of infinitely many affine pieces. Such an observation has clear implications for those wishing to compute such limit metrics numerically. The results of this section have been submitted for publication.

In section 3.5 we show that the previous results can be used to study the effective behaviour of solutions to Newton's second law using the Maupertuis principle. We study in depth the case of one dimensional dynamics and discuss the higher dimensional problem. In particular we are able to provide a rigourous study of motion in highly discontinuous potentials, which seems to be the first study of such problems since [LPV88].

3.2 Homogenisation of Finsler Metrics

In this section we look at the homogenisation of F_ε under standard uniform growth conditions. In particular we determine that F_ε does indeed Γ-converge as $\varepsilon \to 0$ and show that this is equivalent to the induced metrics converging to a norm.

3.2.1 Problem Set-Up

Let $f\colon \mathbb{R}^d \times \mathbb{R}^d \to [0,\infty)$ be a Carathéodory function that is convex and absolutely $1-$homogeneous in the second variable. Suppose also that

$$\alpha\|\xi\| \leq f(s,\xi) \leq \beta\|\xi\|, \quad \forall s,\xi \in \mathbb{R}^d. \tag{3.14}$$

Furthermore suppose that $f(\cdot,\xi)$ is $[0,1]^d$-periodic for all ξ, then set

$$F_\varepsilon(u) := \int_0^1 f\left(\frac{u(\tau)}{\varepsilon}, u'(\tau)\right) \mathrm{d}\tau, \quad u \in W^{1,1}(0,1). \tag{3.15}$$

Let $\xi_1, \xi_2 \in \mathbb{R}^d$ be fixed then define

$$d_\varepsilon(\xi_1,\xi_2) := \inf_{w \in W^{1,\infty}(0,1)} \inf\left\{F_\varepsilon(w)\colon w(0) = \xi_1, w(1) = \xi_2\right\}. \tag{3.16}$$

It follows that d_ε is a metric [BBI01, Chapter 2], by (3.14) it also holds that

$$\alpha\|\xi_1 - \xi_2\| \leq d_\varepsilon(\xi_1,\xi_2) \leq \beta\|\xi_1 - \xi_2\|. \tag{3.17}$$

We will later see that the infimum in (3.16) is infact a minimum.

The main purpose of this section is to determine the Γ-limit of the sequence $(F_\varepsilon)_{\varepsilon>0}$. There are two ways to approach this problem. The first is that of [AV98], and to compute the Γ-limit using direct methods. The second, and the one we take here, is to use the theory of [BPF01]. First we need a definition,

Definition 3.3. *The induced length functional corresponding to a metric d is*

$$L(u) = \int_0^1 f_d(u(\tau), u'(\tau))\, d\tau,$$

where

$$f_d(s,\xi) := \limsup_{t \to 0^+} \frac{d(s, s+t\xi)}{t}. \tag{3.18}$$

Suppose that we prescribe a length functional, as we do here, and its corresponding

induced metric. Does the formula in definition (3.3) recover the original length functional? The answer is yes, as shown in [BBI01, Theorem 2.4.3].

The following theorem is the one we wish to apply to determine the Γ-limit of F_ε.

Theorem 3.4. *It holds that a sequence of metrics $d_\varepsilon\colon \mathbb{R}^d \times \mathbb{R}^d$ satisfying*

$$\alpha\|\xi_1 - \xi_2\| \leq d_\varepsilon(\xi_1, \xi_2) \leq \beta\|\xi_1 - \xi_2\|$$

converges locally uniformly to a metric d if and only if

$$\Gamma - \lim L_\varepsilon(u) = L(u),\ u \in W^{1,\infty}(0,1),$$

where L_ε and L are the induced length functionals for d_ε and d respectively.

Proof. Cf. [BPF01, Theorem 3.1]. □

An immediate corollary of the argument to prove theorem 3.4 is the following result. Define the space

$$\mathcal{A}(\xi_1, \xi_2) := \left\{u \in W^{1,\infty}(0,1)\colon u(0) = \xi_1, u(1) = \xi_2\right\}. \tag{3.19}$$

Corollary 3.5. *It holds that if a sequence of metrics $d_\varepsilon\colon \mathbb{R}^d \times \mathbb{R}^d$ satisfying*

$$\alpha\|\xi_1 - \xi_2\| \leq d_\varepsilon(\xi_1, \xi_2) \leq \beta\|\xi_1 - \xi_2\|$$

converges locally uniformly to a metric d then

$$\Gamma - \lim L_\varepsilon(u) = L(u),\ u \in \mathcal{A}(\xi_1, \xi_2),$$

where L_ε and L are the induced length functionals for d_ε and d respectively.

Proof. Either the argument of [BPF01, Theorem 3.1] can be repeated verbatim with the observation that using curves with prescribed end points has no effect on the result. Alternatively, a proof of the result, using theorem 3.34, can be found in [BD98, Proposition 11.7]. □

We will also need the following property of the induced length functional later. Recall that a function f is positively $1-$homogeneous when $f(\lambda x) = \lambda f(x)$ for all $x \in \mathrm{dom}(f)$ and $\lambda > 0$.

Corollary 3.6. *The function f_d in* (3.18) *in is convex and positively $1-$homogenenous.*

Proof. Cf. [Ven91]. For a detailed exposition of this work consult [Dav04, Chapter 1]. □

So our aim, to prove the Γ-convergence of F_ε, can be achieved by showing that the sequence of induced metrics d_ε converge locally uniformly on $\mathbb{R}^d$. The following lemma makes this task easier.

Lemma 3.7. *If d_ε converge pointwise to d on $\mathbb{R}^d$ then for $K \subset\subset \mathbb{R}^d \times \mathbb{R}^d$ it holds that*

$$\sup_{\xi_1,\xi_2 \in K} |d_\varepsilon(\xi_1,\xi_2) - d(\xi_1,\xi_2)| \to 0$$

as $\varepsilon \to 0$.

Proof. Cf. [BPF01, Proposition 2.3]. □

Therefore in order to prove that the induced metrics converge locally uniformly, it suffices to show that they converge pointwise. What we do in this work is to show something stronger. For our length functionals, which rapidly oscillate in the spatial variable, we show that the oscillations average out in the limit as $\varepsilon \to 0$ and this leads to a translation invariant limit metric. With further argumentation we are able to show that the limit is in fact a norm on $\mathbb{R}^d$. Such a specific analysis of the homogenisation of Finsler metrics in the context of [BPF01] has not been performed to the best of the author's knowledge. Before proceeding with the proof of the pointwise convergence of d_ε let us collect some useful properties of the Finsler functionals F_ε and the induced metric d_ε.

3.2.2 Properties of Finsler Functionals

The following results are basic results in Finsler geometry that we include here for completeness. The following result provides a topological criterion on d_ε for the existence of minimising curves joining two points in F_ε.

Theorem 3.8 (Hopf-Rinow)**.** *If the metric d_ε is complete then every pair of points in $\mathbb{R}^d$ can be joined by a Lipschitz minimising geodesic.*

Proof. Cf. [BBI01, Theorem 2.5.28]. □

The following corollary verifies the topological criterion on d_ε to provide a useful existence result.

Corollary 3.9. *Given any two points in $\mathbb{R}^d$ there exists a Lipschitz minimiser of F_ε joining them.*

Proof. By (3.17) the topology of d_ε is uniformly equivalent to the Euclidean topology on $\mathbb{R}^d$ and is therefore complete. □

The following lemma is known as the 'invariance of length' property for curves. It states that the length of a curve does not depend on the parameterisation.

Lemma 3.10. *Let $u_1\colon I_1 \to \mathbb{R}^d$ and $u_2\colon I_2 \to \mathbb{R}^d$ be such that there exists a map $\sigma\colon I_1 \to I_2$ such that $\sigma' > 0$ and $u_1(\tau) = u_2(\sigma(\tau))$ then*

$$\int_{I_1} f\left(\frac{u_1(\tau)}{\varepsilon}, u_1'(\tau)\right) d\tau = \int_{I_2} f\left(\frac{u_2(\sigma)}{\varepsilon}, u_2'(\sigma)\right) d\sigma. \tag{3.20}$$

Proof. This follows immediately by change of variables and absolute $1-$homogeneity of f in the second variable. See [Jos05, Chapter 1] for a proof of the Riemannian case. □

The following property of the induced metric is also useful when considering a family of metrics. It states that all metrics satisfying $d(\xi_1, \xi_2) \leq \beta\|\xi_1 - \xi_2\|$ are equicontinuous.

Lemma 3.11. *Suppose a metric $d\colon \mathbb{R}^{2d} \to [0, \infty)$ satisfies the bounds of* (3.17)*, then*

$$|d(w, x) - d(y, z)| \leq \beta\left(\|w - y\| + \|x - z\|\right). \tag{3.21}$$

Proof. First observe that

$$|d(w, x) - d(y, z)| \leq |d(w, x) - d(y, x)| + |d(y, x) - d(y, z)|. \tag{3.22}$$

By the triangle inequality we have

$$d(w, x) - d(y, x) \leq d(w, y) \text{ and } -d(y, w) \leq d(w, x) - d(y, x),$$

therefore by (3.17)

$$|d(w, x) - d(y, x)| \leq \beta\|w - y\|. \tag{3.23}$$

Analogously it also holds that $|d(y, x) - d(y, z)| \leq \beta\|x - z\|$, which combined with (3.22) and (3.23) gives the result. □

3.2.3 The Induced Metrics Converge Locally Uniformly

Throughout this subsection we will study F_ε, as defined in subsection 3.2.1. We will need the following definition.

Definition 3.12. *A set $T \subset \mathbb{R}^d$ is relatively dense if there exists $L > 0$ such that $T + [0, L)^d = \mathbb{R}^d$.*

Since f is $[0,1]^d$-periodic in the first variable the set

$$T_\eta^0 := \left\{ \tau \in \mathbb{R}^d \colon |f(s+\tau,\xi) - f(s,\xi)| < \eta(1+\|\xi\|) \right\}, \tag{3.24}$$

is relatively dense in $\mathbb{R}^d$. In particular, owing to the periodicity of f, $\mathbb{Z}^d \subset T_\eta^0$ and hence taking $L = 1$ is sufficient. The following lemma states that the metrics converge pointwise to a translation invariant function.

Lemma 3.13. *For all $\xi \in \mathbb{R}^d$ the limit*

$$\psi(\xi) := \lim_{\varepsilon \to 0} \min_{w \in W^{1,\infty}(0,1)} \{F_\varepsilon(w) \colon w(0) = 0, w(1) = \xi\} \tag{3.25}$$

exists. For $\xi_1, \xi_2 \in \mathbb{R}^d$ the limit,

$$\lim_{\varepsilon \to 0} \min_{w \in W^{1,\infty}(0,1)} \{F_\varepsilon(w) \colon w(0) = \xi_1, w(1) = \xi_2\}$$

exists. Furthermore it holds that

$$\psi(\xi_2 - \xi_1) = \lim_{\varepsilon \to 0} \min_{w \in W^{1,\infty}(0,1)} \{F_\varepsilon(w) \colon w(0) = \xi_1, w(1) = \xi_2\}. \tag{3.26}$$

Proof. The proof that (3.25) exists is identical to the proof of [BD98, Proposition 15.5]. To prove (3.26) let $(\varepsilon_k)_{k=1}^\infty \subset (0,\infty)$ converge to zero. Fix $\xi_1, \xi_2 \in \mathbb{R}^d$. We will show that

$$|d_{\varepsilon_k}(0, \xi_2 - \xi_1) - d_{\varepsilon_k}(\xi_1, \xi_2)| \to 0 \tag{3.27}$$

as $k \to \infty$ and hence (3.26). Fix $\eta > 0$ and let $\xi_1^{\varepsilon_k} \to \xi_1$ as $k \to \infty$ be such that $\tau_k = \xi_1^{\varepsilon_k}/\varepsilon_k \in T_\eta^0$. Such a sequence exists as in the proof of [BD98, Proposition 15.4]. Then

$$\begin{aligned} |d_{\varepsilon_k}(\xi_1, \xi_2) - d_{\varepsilon_k}(0, \xi_2 - \xi_1)| \leq{}& |d_{\varepsilon_k}(\xi_1, \xi_2) - d_{\varepsilon_k}(\xi_1^{\varepsilon_k}, \xi_2 + \xi_1^{\varepsilon_k} - \xi_1)| \\ &+ |d_{\varepsilon_k}(\xi_1^{\varepsilon_k}, \xi_2 + \xi_1^{\varepsilon_k} - \xi_1) - d_{\varepsilon_k}(0, \xi_2 - \xi_1)|. \end{aligned} \tag{3.28}$$

By lemma 3.11 it holds that

$$|d_{\varepsilon_k}(\xi_1, \xi_2) - d_{\varepsilon_k}(\xi_1^{\varepsilon_k}, \xi_2 + \xi_1^{\varepsilon_k} - \xi_1)| \leq 2\beta \|\xi_1 - \xi_1^{\varepsilon_k}\|. \tag{3.29}$$

It remains the analyse the term $d_{\varepsilon_k}(\xi_1^{\varepsilon_k}, \xi_2 + \xi_1^{\varepsilon_k} - \xi_1)$. For each k choose $w_{\varepsilon_k} \in \mathcal{A}(0, \xi_2 - \xi_1)$, the space of Lipschitz curves joining 0 to $\xi_2 - \xi_1$, such that

$$F_{\varepsilon_k}(w_{\varepsilon_k}) \leq d_{\varepsilon_k}(0, \xi_2 - \xi_1) + \varepsilon_k.$$

By our choice of $\xi_1^{\varepsilon_k}$, and using the fact that (3.24) is relatively dense, we obtain

$$\begin{aligned} d_{\varepsilon_k}(\xi_1^{\varepsilon_k}, \xi_2 + \xi_1^{\varepsilon_k} - \xi_1) &\leq F_{\varepsilon_k}(\xi_1^{\varepsilon_k} + w_\varepsilon) \\ &\leq F_{\varepsilon_k}(w_{\varepsilon_k}) + \eta\left(1 + \|w'_{\varepsilon_k}\|_{L^1(0,1)}\right) \\ &\leq d_{\varepsilon_k}(0, \xi_2 - \xi_1) + \varepsilon_k + \eta\left(1 + \|w'_{\varepsilon_k}\|_{L^1(0,1)}\right). \end{aligned} \tag{3.30}$$

It remains to show that $\{\|w'_{\varepsilon_k}\|_{L^1(0,1)}\}_{k=1}^{\infty}$ is bounded. By (3.14), (3.17) and the definition of F_{ε_k} it holds that

$$\alpha\|w'_{\varepsilon_k}\|_{L^1(0,1)} \leq F_{\varepsilon_k}(w_{\varepsilon_k}) \leq d_{\varepsilon_k}(0, \xi_2 - \xi_1) + \varepsilon_k \leq \beta\|\xi_2 - \xi_1\| + \varepsilon_k. \tag{3.31}$$

Combining (3.30) and (3.31) and taking the limit as $k \to \infty$,

$$\begin{aligned} \limsup_{k\to\infty} d_{\varepsilon_k}(\xi_1^{\varepsilon_k}, \xi_2 + \xi_1^{\varepsilon_k} - \xi_1) \leq \lim_{k\to\infty} d_{\varepsilon_k}(0, \xi_2 - \xi_1) \\ + \eta\left(1 + \left(1 + \frac{\beta}{\alpha}\right)\|\xi_2 - \xi_1\|\right). \end{aligned}$$

Since η is arbitrary, it holds that

$$\limsup_{k\to\infty} d_{\varepsilon_k}(\xi_1^{\varepsilon_k}, \xi_2 + \xi_1^{\varepsilon_k} - \xi_1) \leq \lim_{k\to\infty} d_{\varepsilon_k}(0, \xi_2 - \xi_1).$$

The proof that

$$\lim_{k\to\infty} d_{\varepsilon_k}(0, \xi_2 - \xi_1) \leq \liminf_{k\to\infty} d_{\varepsilon_k}(\xi_1^{\varepsilon_k}, \xi_2 + \xi_1^{\varepsilon_k} - \xi_1) \tag{3.32}$$

is achieved in a similar way and hence the right hand side of (3.28) converges to zero. □

Corollary 3.14. *Let ψ be as defined in lemma 3.13. For $\xi_1, \xi_2, \xi_3 \in \mathbb{R}^d$ it holds that*

$$\psi(\xi_1 - \xi_3) \leq \psi(\xi_1 - \xi_2) + \psi(\xi_2 - \xi_3). \tag{3.33}$$

Proof. For each $\varepsilon > 0$ it holds that

$$d_\varepsilon(\xi_3, \xi_1) \leq d_\varepsilon(\xi_3, \xi_2) + d_\varepsilon(\xi_2, \xi_1),$$

taking the limit as $\varepsilon \to 0$ and applying lemma 3.13 gives the result. □

Corollary 3.15. *The function ψ satisfies*

$$\alpha\|\xi\| \leq \psi(\xi) \leq \beta\|\xi\|.$$

Furthermore, it holds that $\psi(\xi) \geq 0$ *for all* $\xi \in \mathbb{R}^d$ *with equality if and only if* $\xi = 0$.

Proof. Let $\xi \in \mathbb{R}^d \setminus \{0\}$. Then, for $\varepsilon > 0$, by (3.17)

$$\alpha\|\xi\| \leq d_\varepsilon(0, \xi) \leq \beta\|\xi\| \tag{3.34}$$

taking the limit as $\varepsilon \to 0$ gives the required inequality. The fact that ψ is positive definite follows immediately from this inequality. □

3.2.4 Γ-convergence of the Length Functionals and Boundary Value Problem

As before, we will study F_ε as defined in subsection 3.2.1. It is not immediately clear that ψ is $1-$homogeneous, and therefore a norm. However, since d_ε is symmetric it follows that $\psi(\xi_1 - \xi_2) = \psi(\xi_2 - \xi_1)$ and therefore by corollaries 3.14 and 3.15 is a metric. In fact, owing to the particular structure extracted in lemma 3.13, it is a translation invariant metric. Before proceeding with the proof of Γ-convergence let us improve the convergence result of lemma 3.13.

Theorem 3.16. *The sequence of functionals* F_ε Γ*-converge with respect to the norm topology of* $L^\infty(0,1)$ *to*

$$\int_0^1 \psi(u'(\tau))\, d\tau,$$

furthermore ψ*, as defined in lemma 3.13, is a norm.*

Proof. By lemmas 3.13 and 3.7 the induced metrics converge locally uniformly to ψ. Recall that ψ is a norm that is uniformly equivalent to the Euclidean norm. Therefore, it follows by theorem 3.4 that the functions F_ε Γ-converge to another functional F that is the length functional induced by ψ. That is,

$$F(u) = \int_0^1 f_\psi(u')\mathrm{d}\tau, \tag{3.35}$$

where f_ψ is defined in corollary 3.6. The next step is to show that $f_\psi(\xi) = \psi(\xi)$ for all $\xi \in \mathbb{R}^d$. To do this observe first that

$$\min\left\{\int_0^1 f_\psi(w'(\tau))\mathrm{d}\tau \colon w \in \mathcal{A}(0,\xi)\right\} \leq \int_0^1 f_\psi(\xi)\mathrm{d}\tau = f_\psi(\xi).$$

Conversely, let $w \in \mathcal{A}(0,\xi)$ be arbitrary, then

$$f_\psi(\xi) = f_\psi\left(\int_0^1 w'(\tau)\mathrm{d}\tau\right) \leq \int_0^1 f_\psi(w'(\tau))\mathrm{d}\tau,$$

by Jensen's inequality. Taking the minimum over all $w \in \mathcal{A}(0,\xi)$ gives that

$$\min\left\{\int_0^1 f_\psi(w'(\tau))\mathrm{d}\tau \colon w \in \mathcal{A}(0,\xi)\right\} = f_\psi(\xi). \tag{3.36}$$

Since the functionals F_ε Γ-converge on the set $\mathcal{A}(0,\xi)$, by corollary 3.5; it follows that

$$\begin{aligned}\min\left\{\int_0^1 f_\psi(w'(\tau))\mathrm{d}\tau \colon w \in \mathcal{A}(0,\xi)\right\} &= \lim_{\varepsilon\to 0}\min\left\{F_\varepsilon(w)\colon w\in\mathcal{A}(0,\xi)\right\}\\ &= \psi(\xi),\end{aligned}$$

and hence $f_\psi \equiv \psi$. □

We collect what information is available to us in the boundary value problem in the theorem below.

Theorem 3.17 (Boundary Value Problem for the Finsler Metric)**.** *Let $\xi_1, \xi_2 \in \mathbb{R}^d$. The functionals F_ε Γ-converge, on $\mathcal{A}(\xi_1,\xi_2)$ with respect to the $L^\infty(0,1)$ topology, to a functional of the form*

$$F(u) := \int_0^1 \psi(u'(\tau))\,d\tau,$$

where ψ, as defined in lemma 3.13, defines a norm on $\mathbb{R}^d$. Furthermore, a function u minimises F, if and only if

$$\int_0^1 \psi(u'(\tau))\,d\tau = \psi\left(\int_0^1 u'(\tau)\,d\tau\right).$$

The functional F always has a minimiser in $\mathcal{A}(\xi_1,\xi_2)$ which is $\bar{u}(\tau) = (\xi_2-\xi_1)\tau + \xi_1$.

Proof. By lemmas 3.13 and 3.7 the induced metrics converge locally uniformly to ψ hence applying corollary 3.5 gives the required Γ-convergence on $\mathcal{A}(\xi_1,\xi_2)$. By Jensen's inequality it holds that for all $u \in \mathcal{A}(\xi_1,\xi_2)$

$$\psi\left(\xi_2-\xi_1\right) = \psi\left(\int_0^1 u'(\tau)\mathrm{d}\tau\right) \le \int_0^1 \psi(u'(\tau))\mathrm{d}\tau,$$

hence $\psi\left(\xi_2-\xi_1\right)$ is a lower bound for the limit functional. This bound is sharp since Jensen's inequality becomes an equality for the function $\bar{u}(\tau) = (\xi_2-\xi_1)\tau+\xi_1$ and hence the minimum value of the limit function is $\psi\left(\xi_2-\xi_1\right)$ attained by $\bar{u}$. The result follows. □

By $1-$homogeneity it follows that ψ can never be strictly (e.g. $\psi(x) = \psi(\frac{1}{2}x)+\psi(\frac{1}{2}x)$), so one cannot decide a priori whether or not $\bar{u}$ is the only limit curve.

3.3 Unbounded Two-Phase Riemannian Metrics

In the following section we consider two phase Riemannian length densities with the $[0,1]^d$-periodic metric coefficient a_ε. The novel feature of this work is that we allow the maximum value of a_ε to be ε-dependant, therefore the oscillation of a_ε will scale with the scaling of the period cell. Our arguments also apply to the case when $\text{Image}(a_\varepsilon) = \{1, \infty\}$.

In the optical interpretation, we design a material whose refractive index is related to the periodicity. In the dynamical interpretation, we will be considering the motion of a particle moving in a discontinuous potential, with the potential wells becoming increasingly deeper as $\varepsilon \to 0$, or infinitely deep. The study of such dynamics has previously been proposed in [LPV88].

Examples of the metric problems have been considered in [BPF01], but so far no systematic study of the general problem has been made. This section aims to address this gap. What is surprising is that while the sequence F_ε always Γ-converges, the corresponding boundary value problem only Γ-converges if a_ε grows sufficiently slowly in ε. As a consequence of the ideas in [BPF01] we are also able to show that the length functionals in the boundary value problem Γ-converge if and only if the corresponding metrics converge locally uniformly to a metric, which is in turn a norm by the previous section.

Before proceeding with the homogenisation study we need to introduce a geometrical concept that will make our subsequent calculations possible.

3.3.1 High Contrast Coefficients

We will say that a set $\Omega_g \subsetneq [0,1]^d$ is an *admissible set* if it is path connected, open, has a Lipschitz boundary and has the property that $\Omega_w := \mathbb{R}^d \setminus \left(\Omega_g + \mathbb{Z}^d\right)$ is path connected. These will be standing assumptions throughout this section. An illustration of this set up is given in figure 3-1. Let us make the following definition.

Definition 3.18 (High Contrast Coefficient)**.** *An admissible set Ω_g has a high contrast coefficient β_0 if for all $x, y \in \partial\Omega_g$ and $\beta > \beta_0$,*

$$\inf_{u\in\mathcal{A}(x,y)} \left\{ \int_0^1 \|u'(\tau)\|\, d\tau \colon u(\tau) \in \partial\Omega_g \ \forall \tau \in (0,1) \right\} < \inf_{u\in\mathcal{A}(x,y)} \left\{ \int_0^1 \beta\|u'(\tau)\|\, d\tau \colon u(\tau) \in \Omega_g \ \forall \tau \in (0,1) \right\}. \quad (3.37)$$

Recall that $\mathcal{A}(x,y)$ is the space of Lipschitz curves joining x to y.

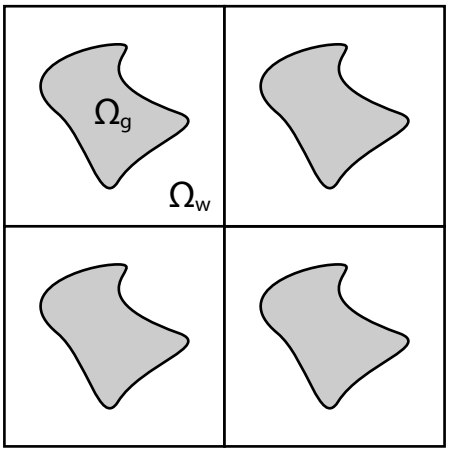

Figure 3-1: *Illustration of a two-phase Riemannian metric.*

The following lemmas give sufficient conditions on Ω_g that ensure the existence of a high contrast coefficient.

Lemma 3.19. *Let Ω_g be an admissible set. If $\partial\Omega_g$ is compact, path connected and differentiable then there exists a high contrast coefficient β_0 for Ω_g.*

Proof. By assumption $\partial\Omega_g$ is a differentiable manifold, therefore it can be equipped with a Riemannian metric [Jos05, Theorem 1.4.1]. The manifold has the distance function

$$d_{\partial\Omega_g}(x,y) := \inf_{u\in\mathcal{A}(x,y)} \left\{ \int_0^1 \|u'(\tau)\| \mathrm{d}\tau \colon u(\tau) \in \partial\Omega_g \ \forall \tau \in (0,1) \right\}, \tag{3.38}$$

for $x, y \in \partial\Omega_g$ [Jos05, Lemma 1.4.1]. The distance function is locally Lipschitz equivalent to the Euclidean distance [Jos05, Corollary 1.4.1] and since $\partial\Omega_g$ is compact there exists constants $\lambda \geq \mu > 0$ such that,

$$\mu\|x-y\| \leq d_{\partial\Omega_g}(x,y) \leq \lambda\|x-y\|, \tag{3.39}$$

for $x, y \in \partial\Omega_g$. Next make the following trivial estimate

$$\inf_{u\in\mathcal{A}(x,y)} \left\{ \int_0^1 \beta\|u'(\tau)\|\mathrm{d}\tau \colon u(\tau) \in \Omega_g \ \forall \tau \in (0,1) \right\} \\ \geq \inf_{u\in\mathcal{A}(x,y)} \left\{ \int_0^1 \beta\|u'(\tau)\|\mathrm{d}\tau \right\} = \beta\|x-y\|. \quad (3.40)$$

Clearly, choosing $\beta > \beta_0 := \lambda$, and combining (3.39) and (3.40) gives (3.37). □

Example 3.20. *If $\Omega_g \subset [0,1]^d$ is a sufficiently small ball, then $\partial\Omega_g$ is a $(d-1)$-sphere and hence Ω_g has a high contrast coefficient.*

The above lemma is sufficient for a very large class of problems, including those that would have physical applications in optics. In the dynamical context such highly discontinuous potentials are not usually considered. However, such a discontinuity would rather serve as a microscopic constraint preventing the particle from entering regions of space. Boundaries of Ω_g with less regularity are still of mathematical interest but are not considered here. In particular it would be an interesting exercise to find a class of Ω_g where (3.37) doesn't hold. We conjecture that if $\partial\Omega_g$ were to have a cusp then (3.37) would fail to hold.

The main reason we need to determine whether or not the high contrast coefficient exists is the following lemma. It states that should the oscillation of a_ε be sufficiently large, then minimising curves do not enter the high contrast regions.

Lemma 3.21. *Let Ω_g be an admissible set with a high contrast coefficient. Then define the function*

$$a(x) := \begin{cases} \beta & \text{if } x \in \Omega_g + \mathbb{Z}^d, \\ 1 & \text{otherwise,} \end{cases} \quad (3.41)$$

for $\beta > \beta_0$, and the Riemannian length functional

$$L(u) := \int_0^1 a(u(\tau))\|u'(\tau)\|\,d\tau. \quad (3.42)$$

Then for any $x, y \in \Omega_w$, let u be a geodesic joining x to y. Then $u(\tau) \in \Omega_w$ for all $\tau \in (0,1)$.

Proof. We prove the contrapositive. Suppose that $\mathrm{Image}(u) \cap \left(\Omega_g + \mathbb{Z}^d\right) \neq \emptyset$. In particular that there exists $\mathbf{x} \in \mathbb{Z}^d$ such that $G(\mathbf{x}) := \mathrm{Image}(u) \cap (\Omega_g + \mathbf{x}) \neq \emptyset$. Set $TG(\mathbf{x}) := \{\tau \in (0,1) \colon u(\tau) \in G(\mathbf{x})\}$, $s = \inf TG(\mathbf{x})$ and $t = \sup TG(\mathbf{x})$.

Then necessarily $s < t$. To see this, observe that by assumption $TG(\mathbf{x}) \neq \emptyset$ therefore there exists $t' \in TG(\mathbf{x})$. By definition $u(t') \in G(\mathbf{x}) \subset \Omega_g + \mathbf{x}$, hence there exists $\rho > 0$

such that $B_\rho(u(t')) \subset \Omega_g + \mathbf{x}$, since the set is open. It follows by continuity that there exists $\delta > 0$ such that $\sigma \in (t' - \delta, t' + \delta)$ implies that $u(\sigma) \in \Omega_g + \mathbf{x}$. By construction $s \leq t' - \delta < t' + \delta \leq t$, hence $s < t$.

The next step is to prove that $u(s), u(t) \in \Omega_w$. We show this for $u(s)$, the argument for $u(t)$ being identical. Suppose that $u(s) \in \Omega_g$, then there exists $\rho > 0$ such that $B_\rho(u(s)) \subset \Omega_g + \mathbf{x}$, again, since the set is open. It follows by continuity that there exists $\delta > 0$ such that $\sigma \in (s-\delta, s+\delta)$ implies that $u(\sigma) \in \Omega_g + \mathbf{x}$. Therefore taking $s' = s - \delta/2$ gives $s' \in (0, 1)$ with $u(s') \in G(\mathbf{x})$ but $s' < s$, a contradiction, hence $u(s) \in \Omega_w$.

Firstly it holds that, since $u(\sigma) \in \Omega_g + \mathbf{x}$ for all $\sigma \in (s, t)$,

$$\int_s^t a(u(\tau))\|u'(\tau)\|\mathrm{d}\tau = \int_s^t \beta\|u'(\tau)\|\mathrm{d}\tau \geq \inf\left\{\int_s^t \beta\|w'(\tau)\|\mathrm{d}\tau \colon w(s) = u(s), w(t) = u(t), w(\tau) \in \Omega_g + \mathbf{x}\ \forall \tau \in (s,t)\right\}.$$

Since $\beta > \beta_0$, where β_0 is the high contrast coefficient, and using the invariance of length under reparameterisations,

$$\int_s^t a(u(\tau))\|u'(\tau)\|\mathrm{d}\tau = \int_s^t \beta\|u'(\tau)\|\mathrm{d}\tau > \inf\left\{\int_s^t \|u'(\tau)\|\mathrm{d}\tau \colon w(s) = u(s), w(t) = u(t), w(\tau) \in \partial\Omega_g + \mathbf{x}\ \forall \tau \in (s,t)\right\}. \quad (3.43)$$

By the Hopf-Rinow theorem [BBI01, Theorem 2.5.28] the above infimum is obtained by a Lipschitz minimising curve $\tilde{u}$. Setting

$$v(\tau) = \begin{cases} \tilde{u}(\tau) & \text{if } \tau \in (s,t) \\ u(\tau) & \text{otherwise,} \end{cases}$$

and using (3.43) gives that $L(u) > L(v)$ and hence u is not geodesic. □

We should also remark that it is not necessary for (3.37) to hold in order to prove lemma 3.21. A similar statement formulated in terms of the convex hull of $\partial\Omega_g$ should broaden the range of shapes that this theory applies to. As already discussed however, the theory here is sufficient for a wide range of physically reasonable examples.

3.3.2 Problem Set-Up

With the notion of a high contrast coefficient we are in a position to rigourously define our unbounded length functionals. Let Ω_g be an admissible set with a high contrast

coefficient β_0 and define the following metric coefficient

$$a_{p,\varepsilon}(x) := \begin{cases} \beta\varepsilon^{-p} & \text{if } x \in \Omega_g + \mathbb{Z}^d, \\ 1 & \text{otherwise,} \end{cases} \tag{3.44}$$

for $\beta > \beta_0$, $p \in (0, \infty)$ and $1 \gg \varepsilon > 0$. We can extend this definition to the case when $p = \infty$ by setting,

$$a_{\infty,\varepsilon}(x) := \begin{cases} +\infty & \text{if } x \in \Omega_g + \mathbb{Z}^d, \\ 1 & \text{otherwise,} \end{cases} \tag{3.45}$$

for $\varepsilon > 0$. With this we define the length functional the we will study, that is

$$F_{p,\varepsilon}(u) := \int_0^1 a_{p,\varepsilon}\left(\frac{u(\tau)}{\varepsilon}\right) \|u'(\tau)\| \, \mathrm{d}\tau. \tag{3.46}$$

The aim of this section is to determine under what conditions $F_{p,\varepsilon}$ Γ-converges. The main feature of our arguments is that we can compare the Γ-convergence of $F_{p,\varepsilon}$ to the Γ-convergence of the functionals

$$F^{\beta}_{\varepsilon}(u) := \int_0^1 a_{\beta}\left(\frac{u(\tau)}{\varepsilon}\right) \|u'(\tau)\| \mathrm{d}\tau \tag{3.47}$$

with the metric coefficient

$$a_{\beta}(x) := \begin{cases} \beta & \text{if } x \in \Omega_g + \mathbb{Z}^d, \\ 1 & \text{otherwise.} \end{cases} \tag{3.48}$$

Recall that by theorem 3.16 it follows that the functionals (3.47) Γ-converge to a functional of the form

$$F^{\beta}_{0}(u) := \int_0^1 \psi(u'(\tau)) \mathrm{d}\tau, \tag{3.49}$$

where ψ is defined in lemma 3.13.

We show that with suitable modifications of the ideas in [BPF01] that we can study the Γ-convergence of $F_{p,\varepsilon}$. This is a natural continuation of the examples at the end of [BPF01]. In order to make these comparisons it is necessary to define a notation corresponding to the induced metrics, to this end define

$$d_{p,\varepsilon}(\xi_1, \xi_2) = \min\left\{F_{p,\varepsilon}(u) \colon u \in W^{1,\infty}(0,1), u(0) = \xi_1, u(1) = \xi_2\right\},$$

it follows it is also a metric [Jos05, Lemma 1.4.1], the minimum exists by the theorem

of Hopf-Rinow [BBI01, Theorem 2.5.28]. It also follows that

$$\|\xi_1 - \xi_2\| \leq d_{p,\varepsilon}(\xi_1, \xi_2) \leq \frac{\beta}{\varepsilon^p}\|\xi_1 - \xi_2\|. \tag{3.50}$$

For clarity of notation in this section we also define

$$d_\varepsilon^\beta(\xi_1, \xi_2) = \min\left\{F_\varepsilon^\beta(u)\colon u \in W^{1,\infty}(0,1), u(0) = \xi_1, u(1) = \xi_2\right\},$$

for more details about these metrics see section 3.2.

3.3.3 Γ-Convergence of the Length Functionals

Let Ω_g be an admissible set with a high contrast coefficient β_0. First we prove that the length functionals $F_{p,\varepsilon}$ Γ-converge on $W^{1,\infty}(0,1)$ with respect to $L^\infty(0,1)$ norm topology for $p \in (0,\infty]$. First we need a technical lemma which states that given $\varepsilon > 0$ and $u \in W^{1,\infty}(0,1)$ there exists a curve no further than $\sqrt{d}\varepsilon$ away from u that doesn't enter the high contrast regions of the Riemannian length density, that is to say precisely that Image$(u) \subset \varepsilon\Omega_w$. We remark that the results of this subsection generalise the examples of [BPF01].

Lemma 3.22. *Let $u \in W^{1,\infty}(0,1)$, then for each $\varepsilon > 0$ there exists $u_\varepsilon^w \in W^{1,\infty}$ such that Image$(u) \subset \varepsilon\Omega_w$ and $\|u - u_\varepsilon^w\|_\infty \leq \sqrt{d}\varepsilon$.*

Proof. Fix $u \in W^{1,\infty}(0,1)$ and $\varepsilon > 0$. Since $\|u\|_\infty < \infty$, it follows that there exist $\mathbf{x}_1, ..., \mathbf{x}_n \in \mathbb{Z}^d$ such that Image$(u) \subset \cup_{i=1}^n \varepsilon([0,1]^d + \mathbf{x}_i)$. Fix $i \in \{1, ..., n\}$ and define $G_i := \text{Image}(u) \cap \varepsilon(\Omega_g + \mathbf{x}_i)$. Let G_i^j be a connected component of G_i; there exists finitely many such connected components since u is Lipschitz. Now fix j. Set $TG_i^j := \{\tau \in (0,1)\colon u(\tau) \in G_i^j\}$, $s_i^j = \inf TG_i^j$ and $t_i^j = \sup TG_i^j$. If $TG_i^j = \emptyset$, then choose a different i, j; if $TG_i^j = \emptyset$ for all i, j then set $u_\varepsilon^w = u$ and we are done. Applying the argument of lemma 3.21, $s_i^j < t_i^j$ and $u(s_i^j), u(t_i^j) \in \partial(\varepsilon(\Omega_g + \mathbf{x}_i))$. Since $\partial\Omega_g$ is path connected there extists a Lipschitz curve joining $u(s_i^j)$ to $u(t_i^j)$ in $\partial(\varepsilon(\Omega_g + \mathbf{x}_i))$ denoted as w_i^j. The fact that w_i^j is Lipschitz continuous follows from the smoothness of $\partial\Omega_g$ [Jos05, Chapters 1 and 8]. Then set

$$u_\varepsilon^w(\tau) := \begin{cases} w_i^j(\tau) & \text{if } \tau \in (s_i^j, t_i^j) \\ u(\tau) & \text{otherwise.} \end{cases}$$

It is clear from the construction that Image$(u) \subset \varepsilon\Omega_w$. Note that, since Ω_g is assumed to be open, Ω_w is closed. It remains to check that $\|u - u_\varepsilon^w\|_\infty \leq \sqrt{d}\varepsilon$. Fix, $\tau \in (s_i, t_i)$

for some $i \in \{1, ..., n\}$, then

$$\|u_\varepsilon^w(\tau) - u(\tau)\| = \|w_i(\tau) - u(\tau)\| \leq \varepsilon \mathrm{diam}(\mathrm{cl}(\Omega_g + \mathbf{x}_i)),$$

since $u(\tau), w_i(\tau) \in \mathrm{cl}(\varepsilon(\Omega_g + \mathbf{x}_i))$ for all $\tau \in (s_i^j, t_i^j)$. Trivially $\mathrm{diam}(\mathrm{cl}(\Omega_g + \mathbf{x}_i)) \leq \sqrt{d}$ and therefore taking the supremum over all τ gives the required estimate. The fact that $u_\varepsilon^w \in W^{1,\infty}$ follows from the regularity of u, w_i and $\partial\Omega_g$. □

The following lemma shows that for the Γ-convergence of F_ε^β, using the previous lemma, it is possible to choose a recovery sequence that never enters the higher contrast region. It is this a key lemma that ensures the Γ-convergence of $F_{p,\varepsilon}$.

Lemma 3.23. *For each $u \in W^{1,\infty}(0,1)$ and each sequence $(\varepsilon_k)_{k=1}^\infty$ converging to 0, there exists a sequence $(u_{\varepsilon_k})_{\varepsilon_k > 0} \subset W^{1,\infty}(0,1)$, converging in $L^\infty(0,1)$ to u, such that*

1. *$\lim_{k\to\infty} F_{\varepsilon_k}^\beta(u_{\varepsilon_k}) = F_0^\beta(u)$, and,*
2. *$\mathrm{Image}(u_{\varepsilon_k}) \subset \varepsilon_k \Omega_w$ for all k.*

Proof. Let $K \subset\subset \mathbb{R}^d$ be such that $\mathrm{Image}(u) \subset \mathrm{int}(K)$. By lemma 3.7 the metrics induced by $F_{\varepsilon_k}^\beta$, denoted $d_{\varepsilon_k}^\beta$, converge locally uniformly to the metric induced by the norm ψ as $k \to \infty$. Therefore, it is possible to choose $(M_k)_{k=1}^\infty \subset \mathbb{N}$ diverging to infinity such that

$$\lim_{k\to\infty} M_k \sup_{\xi_1,\xi_2 \in K} \left| d_{\varepsilon_k}^\beta(\xi_1, \xi_2) - \psi(\xi_2 - \xi_1) \right| = 0.$$

Let $\pi_{M_k} = \{\tau_0, ..., \tau_{M_k}\}$ be a partition of $[0,1]$ such that $|\tau_j - \tau_{j+1}| = 1/M_k$ for $j = 1, ..., M_k$. Define the function u_{ε_k} by

$$u_{\varepsilon_k}(\tau) = u_{\varepsilon_k}^w(\tau) + \mathrm{argmin}_{w \in W_0^{1,\infty}(0,1)} \int_{\tau_{i-1}}^{\tau_i} a_\beta \left(\frac{u(\tau) + w(\tau)}{\varepsilon_k} \right) \|u'(\tau) + w'(\tau)\| \mathrm{d}\tau \quad (3.51)$$

for $\tau \in [\tau_{i-1}, \tau_i]$ in the partition π_{M_k}; the required minimiser exists by the Hopf-Rinow theorem [BBI01, Theorem 2.5.28]. We prove that $u_{\varepsilon_k} \to u$ in $L^\infty(0,1)$. Fix k and $t \in [0,1]$ and suppose that $\tau \in [\tau_{i-1}, \tau_i]$. Then

$$\begin{aligned} \left\| u_{\varepsilon_k}(\tau) - u_{\varepsilon_k}^w(\tau) \right\| &\leq \|u_{\varepsilon_k}(\tau) - u_{\varepsilon_k}(\tau_{i-1})\| + \left\| u_{\varepsilon_k}(\tau_{i-1}) - u_{\varepsilon_k}^w(\tau) \right\|, \\ &= \|u_{\varepsilon_k}(\tau) - u_{\varepsilon_k}(\tau_{i-1})\| + \left\| u_{\varepsilon_k}^w(\tau_{i-1}) - u_{\varepsilon_k}^w(\tau) \right\|, \end{aligned} \quad (3.52)$$

where $u_{\varepsilon_k}^w$ is the sequence from lemma 3.22. By the growth conditions on a_β it holds

that

$$\begin{aligned}\|u_{\varepsilon_k}(\tau) - u_{\varepsilon_k}(\tau_{i-1})\| &\leq d^{\beta}_{\varepsilon_k}(u_{\varepsilon_k}(\tau), u_{\varepsilon_k}(\tau_{i-1})),\\ &\leq d^{\beta}_{\varepsilon_k}(u_{\varepsilon_k}(\tau_i), u_{\varepsilon_k}(\tau_{i-1})),\\ &\leq \beta \left\|u_{\varepsilon_k}(\tau_i) - u_{\varepsilon_k}(\tau_{i-1})\right\|. \end{aligned} \tag{3.53}$$

Since $u^w_{\varepsilon_k} \to u$ uniformly, it follows that the sequence $u^w_{\varepsilon_k}$ is equicontinuous by the converse of the Arzela-Ascoli theorem. Therefore, fix $\eta > 0$ then there exists $\delta > 0$ such that $|x - y| < \delta$ implies that $\|u^w_{\varepsilon_k}(x) - u^w_{\varepsilon_k}(y)\| < \eta$ for all k. Consequently there exists $K \in \mathbb{N}$ such that for $k \geq K$ implies $|\pi_{M_k}| < \delta$, therefore for $k \geq K$, combining (3.52) and (3.53),

$$\left\|u_{\varepsilon_k}(\tau) - u^w_{\varepsilon_k}(\tau)\right\| \leq \beta\eta + \eta.$$

Consequently, since η was arbitrary, $\|u_{\varepsilon_k} - u^w_{\varepsilon_k}\|_\infty \to 0$ as $k \to \infty$, and since $u^w_{\varepsilon_k}$ converges to u in $L^\infty(0,1)$ it holds that $u_{\varepsilon_k} \to u$ in $L^\infty(0,1)$. It remains to show that u_{ε_k} has the desired properties.

$$\begin{aligned}\int_0^1 \psi(u'(\tau))\mathrm{d}\tau &\geq \sum_{i=1}^{M_k} \psi(u(\tau_i) - u(\tau_{i-1}))\\ &\geq \sum_{i=1}^{M_k} d_{p,\varepsilon}(u_{\varepsilon_k}(\tau_i), u_{\varepsilon_k}(\tau_{i-1})) -\\ &\qquad \sum_{i=1}^{M_k} \left|\psi(u(\tau_i) - u(\tau_{i-1})) - d_{p,\varepsilon}(u_{\varepsilon_k}(\tau_i), u_{\varepsilon_k}(\tau_{i-1}))\right|. \end{aligned} \tag{3.54}$$

The first inequality in the above holds since the length functional with density ψ gives rise to an induced metric $d(x,y) = \psi(x-y)$. By construction it is true that

$$\sum_{i=1}^{M_k} d_{p,\varepsilon}(u_{\varepsilon_k}(\tau_i), u_{\varepsilon_k}(\tau_{i-1})) = F_{p,\varepsilon}(u_{\varepsilon_k}),$$

furthermore, since there exists k_0 such that $\mathrm{Image}(u_{\varepsilon_k}) \subset K$ for all $k \geq k_0$,

$$\begin{aligned}&\sum_{i=1}^{M_k} \left|\psi(u(\tau_i) - u(\tau_{i-1})) - d_{p,\varepsilon}(u_{\varepsilon_k}(\tau_i), u_{\varepsilon_k}(\tau_{i-1}))\right|\\ &\qquad\qquad \leq M_k \sup_{\xi_1,\xi_2 \in K} |d_{\varepsilon_k}(\xi_1,\xi_2) - \psi(\xi_2 - \xi_1)|.\end{aligned}$$

Hence by (3.70), the choice of $(M_k)_{k=1}^{\infty}$ and the liminf inequality,

$$\int_0^1 \psi(u'(\tau))\mathrm{d}\tau \geq \limsup_{k\to\infty} F_{p,\varepsilon}(u_{\varepsilon_k}) \geq \int_0^1 \psi(u'(\tau))\mathrm{d}\tau.$$

The last thing we need to show is that $\mathrm{Image}(u_{\varepsilon_k}) \subset \varepsilon_k \Omega_w$ for all k. By the construction, as given in (3.51), u_{ε_k} is constructed as a piecewise geodesic curve joining points along $u^w_{\varepsilon_k}$. By lemma 3.21 these geodesic pieces do not enter $\Omega_g + \mathbb{Z}^d$, hence the result holds. □

Theorem 3.24. *Let Ω_g have a high contrast coefficient β_0 and let $\beta > \beta_0$. Then*

$$\Gamma(L^\infty(0,1)) - \lim F_{p,\varepsilon}(u) = F_0^\beta(u),\ \forall u \in W^{1,\infty}(0,1), \forall p \in (0,\infty].$$

Proof. Fix $p \in (0,\infty]$ and let $(\varepsilon_k)_{k=1}^{\infty}$ be a sequence converging to 0. Let $u_{\varepsilon_k} \to u$ in $L^\infty(0,1)$, then, since $a_{p,\varepsilon_k} \geq a_\beta$ for $\varepsilon_k < 1$,

$$\liminf_{k\to\infty} F_{p,\varepsilon_k}(u_{\varepsilon_k}) \geq \liminf_{k\to\infty} F^\beta_{\varepsilon_k}(u_{\varepsilon_k}) \geq F_0^\beta(u),$$

the second inequality being the liminf inequality for the Γ-convergence of $F^\beta_{\varepsilon_k}$. Applying lemma 3.23 we obtain a sequence u_{ε_k} converging to u in $L^\infty(0,1)$ where $\lim_{k\to\infty} F^\beta_{\varepsilon_k}(u_{\varepsilon_k}) = F_0^\beta(u)$ and $\mathrm{Image}(u_{\varepsilon_k}) \subset \varepsilon_k\Omega_w$ for all k. Since $a_{p,\varepsilon_k} = a_\beta$ on Ω_w it follows that

$$\lim_{k\to\infty} F_{p,\varepsilon_k}(u_{\varepsilon_k}) = \lim_{k\to\infty} F^\beta_{\varepsilon_k}(u_{\varepsilon_k}).$$

Hence the sequence F_{p,ε_k} Γ-converges. □

While this Γ-convergence result is nontrivial, as the sequence of functionals are not bounded uniformly, the usual convergence about the convergence of minimisers can be seen in an easier fashion. Namely, the minimiser of $F_{p,\varepsilon}$ is 0, which is also unique and this clearly converges to the unique minimiser of F_0^β. What is significantly more interesting is to understand the Γ-convergence of $F_{p,\varepsilon}$ on a smaller space, typically the space of Lipschitz curves joining two fixed points. Here, the minimal curves are not necessarily trivial, and to understand their effective behaviour proves to be more challenging in the context of the unbounded length functionals we consider here.

3.3.4 Γ-Convergence of the Boundary Value Problem

As before, let Ω_g be an admissible set with a high contrast coefficient β_0. We have seen in the previous subsection the unbounded length functionals $F_{p,\varepsilon}$ indeed Γ-converge, regardless of the value of p. Little information can be inferred from this statement however, with regards to geodesics joining two points. In this subsection we study the

effective description of geodesics joining two points. The surprising fact we investigate here is that the existence of the Γ-limit of $F_{p,\varepsilon}$ depends on the value of p. In fact we have that the sequence Γ-converges on the space of curves joining any two points if and only if the induced metric converges locally uniformly.

The aim of this subsection is to relate the Γ-convergence of the unbounded length functionals with the local uniform convergence of the induced metrics. The equivalence of the modes of convergence for metrics that are uniformly equivalent to the Euclidean distance was the focus of [BPF01]. Here we show that the induced metrics always converge locally uniformly, should the contrast grow sufficiently slowly. This in turn implies the Γ-convergence of the functionals.

The Induced Metrics Converge Locally Uniformly to a Norm ($p < 1$)

First we have a technical lemma which states that given any point in $\xi \in \mathbb{R}^d$ and $\varepsilon > 0$ there exists a point in $\xi' \in \varepsilon\Omega_w$ such that $\|\xi - \xi'\| \leq \sqrt{d}\varepsilon$.

Lemma 3.25. *For all $\xi \in \mathbb{R}^d, \varepsilon > 0$ there exists $\xi_\varepsilon \in \varepsilon\Omega_w$ be such that $\|\xi_\varepsilon - \xi\| \leq \sqrt{d}\varepsilon$*

Proof. Fix $\varepsilon > 0$, then $\xi \in \varepsilon([0,1]^d + \mathbf{x})$ for some $\mathbf{x} \in \mathbb{Z}^d$. Then take $\xi_\varepsilon \in \varepsilon(\Omega_w + \mathbf{x})$, which exists since $\Omega_g \subsetneq [0,1]^d$. The required estimate follows immediately. □

The following lemma shows that, for $p < 1$, the convergence of the induced metrics is equivalent to the convergence of the same metrics evaluated at nearby points of $\varepsilon\Omega_w$. It will be easier to determine the latter convergence as we have more information about geodesics joining points in $\varepsilon\Omega_w$ due to lemma 3.21.

Lemma 3.26. *Let $p < 1$, $\xi_1, \xi_2 \in \mathbb{R}^d$, $\varepsilon > 0$ and $\xi_{1,\varepsilon}, \xi_{2,\varepsilon} \in \varepsilon\Omega_w$ be such that $\|\xi_{1,\varepsilon} - \xi_1\| \leq \sqrt{d}\varepsilon$, $\|\xi_{2,\varepsilon} - \xi_2\| \leq \sqrt{d}\varepsilon$. The limit*

$$\lim_{\varepsilon\to 0} d_{p,\varepsilon}(\xi_1, \xi_2) = \lim_{\varepsilon\to 0} \min_{u\in\mathcal{A}(\xi_1,\xi_2)} F_{p,\varepsilon}(u) \tag{3.55}$$

exists if and only if the limit

$$\lim_{\varepsilon\to 0} d_{p,\varepsilon}(\xi_{1,\varepsilon}, \xi_{2,\varepsilon}) = \lim_{\varepsilon\to 0} \min_{u\in\mathcal{A}(\xi_{1,\varepsilon},\xi_{2,\varepsilon})} F_{p,\varepsilon}(u) \tag{3.56}$$

exists.

Proof. Fix $\varepsilon \in (0,1)$. Then, using the triangle inequality and (3.50),

$$\begin{aligned} d_{p,\varepsilon}(\xi_1,\xi_2) &\leq d_{p,\varepsilon}(\xi_1,\xi_{1,\varepsilon}) + d_{p,\varepsilon}(\xi_{1,\varepsilon},\xi_{2,\varepsilon}) + d_{p,\varepsilon}(\xi_{2,\varepsilon},\xi_2) \\ &\leq \frac{\beta}{\varepsilon^p}\|\xi_1 - \xi_{1,\varepsilon}\| + d_{p,\varepsilon}(\xi_{1,\varepsilon},\xi_{2,\varepsilon}) + \frac{\beta}{\varepsilon^p}\|\xi_{2,\varepsilon} - \xi_2\| \\ &\leq d_{p,\varepsilon}(\xi_{1,\varepsilon},\xi_{2,\varepsilon}) + 2\beta\sqrt{d}\varepsilon^{1-p}. \end{aligned}$$

Furthermore,

$$\begin{aligned} d_{p,\varepsilon}(\xi_1,\xi_2) &\geq d_{p,\varepsilon}(\xi_{1,\varepsilon},\xi_{2,\varepsilon}) - d_{p,\varepsilon}(\xi_{1,\varepsilon},\xi_1) - d_{p,\varepsilon}(\xi_2,\xi_{2,\varepsilon}) \\ &\geq d_{p,\varepsilon}(\xi_{1,\varepsilon},\xi_{2,\varepsilon}) - \frac{\beta}{\varepsilon^p}\|\xi_{1,\varepsilon} - \xi_1\| - \frac{\beta}{\varepsilon^p}\|\xi_2 - \xi_{2,\varepsilon}\| \\ &\geq d_{p,\varepsilon}(\xi_{1,\varepsilon},\xi_{2,\varepsilon}) - 2\beta\sqrt{d}\varepsilon^{1-p}. \end{aligned}$$

Therefore

$$d_{p,\varepsilon}(\xi_{1,\varepsilon},\xi_{2,\varepsilon}) - 2\beta\sqrt{d}\varepsilon^{1-p} \leq d_{p,\varepsilon}(\xi_1,\xi_2) \leq d_{p,\varepsilon}(\xi_{1,\varepsilon},\xi_{2,\varepsilon}) + 2\beta\sqrt{d}\varepsilon^{1-p}. \tag{3.57}$$

If the limit (3.56) exists then taking the limit as $\varepsilon \to 0$ gives that, by (3.57), the limit (3.55) exists. The converse statement is proved in an identical fashion. □

The next lemma is intended to be used with lemma 3.26. In essence, it states that we may 'unfold' the ε-dependant metric and rather than considering the minimisation of ε-dependant metrics we can consider the minimisation of a single-scale metric joining ε-dependant end points. The benefit of doing this is so that lemma 3.21 can be applied to the unfolded metric and guarantee that our sequence of geodesics all lie in $\varepsilon\Omega_w$ for our chosen endpoints.

Lemma 3.27. *Let $p < 1$, $\xi_1, \xi_2 \in \mathbb{R}^d$, $\varepsilon > 0$ and $\xi_{1,\varepsilon}, \xi_{2,\varepsilon} \in \varepsilon\Omega_w$ be such that $\|\xi_{1,\varepsilon} - \xi_1\| \leq \sqrt{d}\varepsilon$, $\|\xi_{2,\varepsilon} - \xi_2\| \leq \sqrt{d}\varepsilon$. Then*

$$\lim_{\varepsilon\to 0} \min_{u\in\mathcal{A}(\xi_{1,\varepsilon},\xi_{2,\varepsilon})} F_{p,\varepsilon}(u) = \lim_{\varepsilon\to 0} \min_{u\in W^{1,\infty}(0,1)} \left\{ \varepsilon \int_0^1 a_{p,\varepsilon}(u(\tau))\|u'(\tau)\|\,d\tau \colon u(0) = \frac{\xi_{1,\varepsilon}}{\varepsilon}, u(1) = \frac{\xi_{2,\varepsilon}}{\varepsilon} \right\},$$

should either limit exist.

Proof. Fix $\varepsilon > 0$, $u \in W^{1,\infty}(0,1)$ with $u(0) = \xi_{1,\varepsilon}, u(1) = \xi_{2,\varepsilon}$. Then

$$F_{p,\varepsilon}(u) = \int_0^1 a_{p,\varepsilon}\left(\frac{u(\tau)}{\varepsilon}\right)\|u'(\tau)\|\mathrm{d}\tau = \varepsilon\int_0^1 a_{p,\varepsilon}\left(w(\tau)\right)\|w'(\tau)\|\mathrm{d}\tau,$$

where $w = u/\varepsilon$, hence $w(0) = \xi_{1,\varepsilon}/\varepsilon, w(1) = \xi_{2,\varepsilon}/\varepsilon$. Since the correspondence between u and w is one-to-one the result follows by taking the infimum over u, or equivalently w, and then passing to the limit. □

The following lemma connects the unfolded homogenisation problem to the unfolded homogenisation problem for the sequence F_ε^β. The heart of the proof is elementary, for our given endpoints the geodesic remains in $\varepsilon\Omega_w$, for curves in $\varepsilon\Omega_w$ the functionals F_ε^β and $F_{p,\varepsilon}$ give equivalent measures of length.

Lemma 3.28. *Let $p < 1$, $\xi_1, \xi_2 \in \mathbb{R}^d$, $\varepsilon > 0$ and $\xi_{1,\varepsilon}, \xi_{2,\varepsilon} \in \varepsilon\Omega_w$ be such that $\|\xi_{1,\varepsilon} - \xi_1\| \leq \sqrt{d}\varepsilon$, $\|\xi_{2,\varepsilon} - \xi_2\| \leq \sqrt{d}\varepsilon$. . For $1 > \varepsilon > 0$ it follows that*

$$\lim_{\varepsilon\to 0} \min_{u\in W^{1,\infty}(0,1)} \left\{ \varepsilon \int_0^1 a_{p,\varepsilon}(u(\tau))\|u'(\tau)\| d\tau \colon u(0) = \frac{\xi_{1,\varepsilon}}{\varepsilon}, u(1) = \frac{\xi_{2,\varepsilon}}{\varepsilon} \right\} = \lim_{\varepsilon\to 0} \min_{u\in W^{1,\infty}(0,1)} \left\{ \varepsilon \int_0^1 a_{\beta}(u(\tau))\|u'(\tau)\| d\tau \colon u(0) = \frac{\xi_{1,\varepsilon}}{\varepsilon}, u(1) = \frac{\xi_{2,\varepsilon}}{\varepsilon} \right\}.$$

Proof. Fix $\varepsilon \in (0, 1)$. Then,

$$\min_{u\in W^{1,\infty}(0,1)} \left\{ \varepsilon \int_0^1 a_{p,\varepsilon}(u(\tau))\|u'(\tau)\| \mathrm{d}\tau \colon u(0) = \frac{\xi_{1,\varepsilon}}{\varepsilon}, u(1) = \frac{\xi_{2,\varepsilon}}{\varepsilon} \right\} = \min_{u\in W^{1,\infty}(0,1)} \left\{ \varepsilon \int_0^1 a_{p,\varepsilon}(u(\tau))\|u'(\tau)\| \mathrm{d}\tau \colon u(0) = \frac{\xi_{1,\varepsilon}}{\varepsilon}, u(1) = \frac{\xi_{2,\varepsilon}}{\varepsilon}, u(\tau) \in \Omega_w \ \forall \tau \right\} \quad (3.58)$$

by lemma 3.21, observing that by definition $\xi_{1,\varepsilon}/\varepsilon, \xi_{2,\varepsilon}/\varepsilon \in \Omega_w$. Since $a_{p,\varepsilon} = a_\beta$ on Ω_w

$$\min_{u\in W^{1,\infty}(0,1)} \left\{ \varepsilon \int_0^1 a_{p,\varepsilon}(u(\tau))\|u'(\tau)\| \mathrm{d}\tau \colon u(0) = \frac{\xi_{1,\varepsilon}}{\varepsilon}, u(1) = \frac{\xi_{2,\varepsilon}}{\varepsilon}, u(\tau) \in \Omega_w \ \forall \tau \right\} = \min_{u\in W^{1,\infty}(0,1)} \left\{ \varepsilon \int_0^1 a_{\beta}(u(\tau))\|u'(\tau)\| \mathrm{d}\tau \colon u(0) = \frac{\xi_{1,\varepsilon}}{\varepsilon}, u(1) = \frac{\xi_{2,\varepsilon}}{\varepsilon}, u(\tau) \in \Omega_w \ \forall \tau \right\}. \quad (3.59)$$

Applying lemma 3.21 again gives that

$$\min_{u\in W^{1,\infty}(0,1)} \left\{ \varepsilon \int_0^1 a_{\beta}(u(\tau))\|u'(\tau)\| \mathrm{d}\tau \colon u(0) = \frac{\xi_{1,\varepsilon}}{\varepsilon}, u(1) = \frac{\xi_{2,\varepsilon}}{\varepsilon}, u(\tau) \in \Omega_w \ \forall \tau \right\} = \min_{u\in W^{1,\infty}(0,1)} \left\{ \varepsilon \int_0^1 a_{\beta}(u(\tau))\|u'(\tau)\| \mathrm{d}\tau \colon u(0) = \frac{\xi_{1,\varepsilon}}{\varepsilon}, u(1) = \frac{\xi_{2,\varepsilon}}{\varepsilon} \right\}. \quad (3.60)$$

Tracing the chain of equivalences through equations (3.58), (3.59) and (3.60) and then taking the limit as $\varepsilon \to 0$ gives the result. □

The following lemma is a statement of the analogues of lemmas 3.26 and 3.27 for the

sequence F_ε^β.

Lemma 3.29. *Let $p < 1$, $\xi_1, \xi_2 \in \mathbb{R}^d$, $\varepsilon > 0$ and $\xi_{1,\varepsilon}, \xi_{2,\varepsilon} \in \varepsilon\Omega_w$ be such that $\|\xi_{1,\varepsilon} - \xi_1\| \leq \sqrt{d}\varepsilon$, $\|\xi_{2,\varepsilon} - \xi_2\| \leq \sqrt{d}\varepsilon$. Then*

$$\liminf_{\varepsilon\to 0} \left\{ F_\beta(u) \colon W^{1,\infty}(0,1), u(0) = \xi_1, u(1) = \xi_2 \right\} = \liminf_{\varepsilon\to 0} \left\{ \varepsilon \int_0^{1/\varepsilon} a_\beta(u(\tau)) \|u'(\tau)\| \, d\tau \colon u \in W^{1,\infty}(0,1), u(0) = \frac{\xi_{1,\varepsilon}}{\varepsilon}, u(1) = \frac{\xi_{2,\varepsilon}}{\varepsilon} \right\}.$$

Furthermore, the limit

$$\lim_{\varepsilon\to 0} \min_{u \in \mathcal{A}(\xi_1,\xi_2)} F_\varepsilon^\beta(u) \tag{3.61}$$

exists, if and only if, the limit

$$\lim_{\varepsilon\to 0} \min_{u \in \mathcal{A}(\xi_{1,\varepsilon},\xi_{2,\varepsilon})} F_\varepsilon^\beta(u) \tag{3.62}$$

exists.

Proof. The first part follows exactly as in lemma 3.27. The second part is demonstrated using an identical proof to that for lemma 3.26. □

The following theorem combines lemmas 3.26, 3.27 and 3.29 to establish the pointwise convergence of metrics for $p < 1$.

Theorem 3.30. *Let $p < 1$. Then the limit*

$$\lim_{\varepsilon\to 0} d_{p,\varepsilon}(\xi_1, \xi_2) = \lim_{\varepsilon\to 0} \min_{u \in \mathcal{A}(\xi_1,\xi_2)} F_{p,\varepsilon}(u) \tag{3.63}$$

exists for all $\xi_1, \xi_2 \in \mathbb{R}^d$. Furthermore,

$$\lim_{\varepsilon\to 0} d_{p,\varepsilon}(\xi_1, \xi_2) = \lim_{\varepsilon\to 0} \min_{u \in \mathcal{A}(\xi_1,\xi_2)} F_\varepsilon^\beta(u) = \psi(\xi_2 - \xi_1) \tag{3.64}$$

where ψ is defined in lemma 3.13, for any $\beta > \beta_0$.

Proof. By lemma 3.13 the limit

$$\lim_{\varepsilon\to 0} \min_{u \in \mathcal{A}(\xi_1,\xi_2)} F_\varepsilon^\beta(u)$$

exists. Applying lemma 3.25, for each $\varepsilon > 0$, we obtain the existence of $\xi_{1,\varepsilon}, \xi_{2,\varepsilon} \in \varepsilon\Omega_w$ be such that $\|\xi_{1,\varepsilon} - \xi_1\| \leq \sqrt{d}\varepsilon$, $\|\xi_{2,\varepsilon} - \xi_2\| \leq \sqrt{d}\varepsilon$. The result then follows by applying lemmas 3.29, 3.28, 3.26 and then 3.27 in that order. □

The Induced Metrics Fail to Converge Pointwise ($p \geq 1$)

In this section we show that for $p \geq 1$ the limit

$$\lim_{\varepsilon \to 0} d_{p,\varepsilon}(\xi_1, \xi_2) = \lim_{\varepsilon \to 0} \inf_{u \in W^{1\infty}(0,1)} \{F_{p,\varepsilon}(u) \colon u(0) = \xi_1, u(1) = \xi_2\}$$

does not exist, in contrast to the previous section, which showed that for $p < 1$ it does. Consequently, the sequence $F_{p,\varepsilon}$ fails to Γ-converge on $\mathcal{A}(\xi_1, \xi_2)$ for all $\xi_1, \xi_2 \in \mathbb{R}^d$. Since otherwise by theorem 3.2 the above limit would exist, as Γ-convergence implies the convergence of minimum values.

We prove a particular statement that relies on a simple geometric assumption. The principle behind the argument is clear, and it is also clear that such an argument could be generalised to other cases. We do not attempt such a general argument here, but suggest that it would be possible using the fact that $a(\xi\tau)$ is quasiperiodic in τ for fixed ξ.

Theorem 3.31. *Suppose that $\xi_2 := (1/2, 1/2) \in \Omega_g$ and that $\xi_1 := (0,0) \in \Omega_w$. Then the limit*

$$\lim_{\varepsilon \to 0} d_{p,\varepsilon}(\xi_1, \xi_2) = \lim_{\varepsilon \to 0} \inf_{u \in W^{1\infty}(0,1)} \{F_{p,\varepsilon}(u) \colon u(0) = \xi_1, u(1) = \xi_2\}$$

does not exist.

Proof. Set, for $k \in \mathbb{N}$,

$$\tilde{\varepsilon}_k := \frac{1}{2k}.$$

It follows immediately that $\xi_2 \in \tilde{\varepsilon}_k \Omega_w$ for all k. Hence by lemma 3.21 the solution to the problem

$$\min_{u \in W^{1\infty}(0,1)} \{F_{p,\tilde{\varepsilon}_k}(u) \colon u(0) = \xi_1, u(1) = \xi_2\},$$

which we denote by $w_{\tilde{\varepsilon}_k}$, is such that $\mathrm{Image}(w_{\tilde{\varepsilon}_k}) \subset \tilde{\varepsilon}_k \Omega_w$. By theorem 3.30, with $\xi_{1,\tilde{\varepsilon}_k} = \xi_1$ and $\xi_{2,\tilde{\varepsilon}_k} = \xi_2$ for all $\tilde{\varepsilon} > 0$, it holds that

$$\begin{aligned}
&\lim_{k \to \infty} \min_{u \in W^{1\infty}(0,1)} \{F_{p,\tilde{\varepsilon}_k}(u) \colon u(0) = \xi_1, u(1) = \xi_2\} \\
&\quad = \lim_{k \to \infty} \min_{u \in W^{1\infty}(0,1)} \left\{F^{\beta}_{\tilde{\varepsilon}_k}(u) \colon u(0) = \xi_1, u(1) = \xi_2\right\} \\
&\quad = \psi(\xi_2 - \xi_1).
\end{aligned}$$

Now define for $k \in \mathbb{N}$,

$$\varepsilon_k = \frac{1}{2k+1}.$$

Then $\xi_2 \in \varepsilon_k(\Omega_g + \mathbb{Z}^d)$ for all k. Denote the solution to the problem

$$\min_{u \in W^{1\infty}(0,1)} \{F_{p,\varepsilon_k}(u) \colon u(0) = \xi_1, u(1) = \xi_2\},$$

by w_{ε_k}. It follows by lemma 3.21 that for each k there exists $\tau_k \in (0,1)$ such that $w_{\varepsilon_k}(\tau) \in \varepsilon_k \Omega_w$ for $\tau \in [0, \tau_k]$ and $w_{\varepsilon_k}(\tau) \in \varepsilon_k(\Omega_g + (k,k))$ for $\tau \in (\tau_k, 1]$. Since Ω_g is open there exists a $\rho > 0$ such that $B_\rho((1/2,1/2)) \subset \Omega_g$. Hence, for all k, $\varepsilon_k(B_\rho((1/2,1/2)) + (k,k)) \subset \varepsilon_k(\Omega_g + (k,k))$. For each k there exists, by continuity, $\sigma_k \in (\tau_k, 1)$ such that $w_{\varepsilon_k}(\sigma_k) \in \partial\left(\varepsilon_k(B_\rho((1/2,1/2)) + (k,k))\right)$. Therefore,

$$\begin{aligned}
F_{p,\varepsilon_k}(w_{\varepsilon_k}) &= \int_0^{\tau_k} a_{p,\varepsilon}\left(\frac{w_{\varepsilon_k}(\tau)}{\varepsilon_k}\right) \|w'_{\varepsilon_k}(\tau)\| \mathrm{d}\tau + \int_{\tau_k}^1 a_{p,\varepsilon}\left(\frac{w_{\varepsilon_k}(\tau)}{\varepsilon_k}\right) \|w'_{\varepsilon_k}(\tau)\| \mathrm{d}\tau \\
&\geq \int_0^{\tau_k} a_{p,\varepsilon}\left(\frac{w_{\varepsilon_k}(\tau)}{\varepsilon_k}\right) \|w'_{\varepsilon_k}(\tau)\| \mathrm{d}\tau + \int_{\sigma_k}^1 a_{p,\varepsilon}\left(\frac{w_{\varepsilon_k}(\tau)}{\varepsilon_k}\right) \|w'_{\varepsilon_k}(\tau)\| \mathrm{d}\tau \\
&= d_{p,\varepsilon}(\xi_1, w_{\varepsilon_k}(\tau_k)) + d_{p,\varepsilon}(w_{\varepsilon_k}(\sigma_k), \xi_2) =: I_1 + I_2,
\end{aligned}$$

since any geodesic curve is locally geodesic. We now analyse I_1. First observe that by construction $\|w_{\varepsilon_k}(\tau_k) - \xi_2\| \leq \sqrt{d}\varepsilon_k$ and $w_{\varepsilon_k}(\tau_k) \in \varepsilon_k \Omega_w$. Setting $\xi_{1,\varepsilon_k} = \xi_1$ and $\xi_{2,\varepsilon_k} = w_{\varepsilon_k}(\tau_k)$ for all k, then applying theorem 3.30 it follows that

$$\lim_{k\to\infty} d_{p,\varepsilon}(\xi_1, w_{\varepsilon_k}(\tau_k)) = \lim_{k\to\infty} d^\beta_{\varepsilon_k}(\xi_1, \xi_2) = \psi(\xi_2 - \xi_1),$$

It remains to estimate I_2, as $d_{p,\varepsilon}(w_{\varepsilon_k}(\sigma_k), \xi_2)$ is the distance between the centre of the ball $\varepsilon_k(B_\rho((1/2,1/2)) + (k,k))$ and its boundary. Since the ball is contained in Ω_g it follows that

$$\begin{aligned}
d_{p,\varepsilon}(w_{\varepsilon_k}(\sigma_k), \xi_2) &= \frac{\beta}{\varepsilon_k^p} \|w_{\varepsilon_k}(\sigma_k) - \xi_2\| \\
&= \beta \varepsilon_k^{1-p} \rho.
\end{aligned}$$

Therefore, combining these results and taking the limit gives that

$$\lim_{k\to\infty} F_{p,\varepsilon_k}(w_{\varepsilon_k}) \geq \begin{cases} \infty & \text{if } p > 1, \\ \psi(\xi_2 - \xi_1) + \beta\rho & \text{if } p = 1. \end{cases}$$

Hence there exists sequences $(\tilde{\varepsilon}_k)_{k=1}^\infty$ and $(\varepsilon_k)_{k=1}^\infty$ converging to 0 such that

$$d_{p,\tilde{\varepsilon}_k}(\xi_1, \xi_2) \geq d_{p,\varepsilon_k}(\xi_1, \xi_2) + \beta\rho,$$

thus the result is shown. □

The geometrical assumption is not particularly restrictive. If we assumed that $a_{p,\varepsilon}$ defines the refractive index of an optical medium and assuming that Ω_g is an inclusion of high contrast material in an otherwise homogeneous background. Then the assumption in theorem 3.31 would apply to the case when the inclusion is contained in the interior of the unit cell and the inclusion is at the centre of the unit cell.

The Equivalence of Γ-convergence and Metric Convergence ($p < 1$)

In this section we prove a stronger statement than that of the length functionals Γ-converging for the boundary value problem. We show that the boundary value problem Γ-converges if and only if the induced metrics converge locally uniformly. This extends the theory of [BPF01] to the case of unbounded two-phase Riemannian length functionals. The following lemma shows that we can improve the bounds on the induced metric so that d_ε is almost uniformly equivalent to the Euclidean metric.

Lemma 3.32. *Let $p < 1$, $\varepsilon > 0$ and $\xi_1, \xi_2 \in \mathbb{R}^d$. Then there exists $C_1, C_2 > 0$ such that*

$$\|\xi_1 - \xi_2\| - C_1\varepsilon \leq d_{p,\varepsilon}(\xi_1, \xi_2) \leq \beta\|\xi_1 - \xi_2\| + C_2\varepsilon^{1-p}$$

Proof. Let u be the geodesic joining ξ_1 to ξ_2. By lemma 3.21 it follows that the set $T := \{\tau \in (0,1) \colon u(\tau) \in \varepsilon\left(\Omega_g + \mathbb{Z}^d\right)\}$ takes one of the following forms.

$$\emptyset \text{ or } [0, \tau_1) \text{ or } (\tau_2, 1] \text{ or } [0, \tau_1) \cup (\tau_2, 1] \text{ or } [0, 1]$$

for some $\tau_1, \tau_2 \in (0,1)$ with $\tau_1 < \tau_2$. Suppose first that $T = [0, \tau_1) \cup (\tau_2, 1]$; the cases when $\tau_2 = 1$ or $\tau_1 = 0$ following in an identical fashion. Observe that

$$\begin{aligned} F_{p,\varepsilon}(u) &= \int_0^{\tau_1} a_{p,\varepsilon}\left(\frac{u(\tau)}{\varepsilon}\right)\|u'(\tau)\|\mathrm{d}\tau + \int_{\tau_1}^{\tau_2} a_{p,\varepsilon}\left(\frac{u(\tau)}{\varepsilon}\right)\|u'(\tau)\|\mathrm{d}\tau \\ &\quad + \int_{\tau_2}^{1} a_{p,\varepsilon}\left(\frac{u(\tau)}{\varepsilon}\right)\|u'(\tau)\|\mathrm{d}\tau \\ &= d_{p,\varepsilon}(\xi_1, u(\tau_1)) + d_{p,\varepsilon}(u(\tau_1), u(\tau_2)) + d_{p,\varepsilon}(u(\tau_2), \xi_2). \end{aligned} \tag{3.65}$$

Observe that by construction $\|\xi_1 - u(\tau_1)\| \leq \sqrt{d}\varepsilon$ and $\|\xi_2 - u(\tau_2)\| \leq \sqrt{d}\varepsilon$, and hence by the growth condition (3.50) it follows that

$$d_{p,\varepsilon}(\xi_1, u(\tau_1)) \leq \beta\sqrt{d}\varepsilon^{1-p} \text{ and } d_{p,\varepsilon}(u(\tau_2), \xi_2) \leq \beta\sqrt{d}\varepsilon^{1-p}.$$

Using lemma 3.21 and the fact that $u(\tau_1), u(\tau_2) \in \varepsilon\Omega_w$, it follows that $a_{p,\varepsilon}(u(\tau)) =$

$1 = a_\beta(u(\tau))$ for all $\tau \in [\tau_1, \tau_2]$ and hence $d_{p,\varepsilon}(u(\tau_1), u(\tau_2)) = d^\beta_\varepsilon(u(\tau_1), u(\tau_2))$. By the triangle inequality and (3.50) we have

$$\begin{aligned} d^\beta_\varepsilon(u(\tau_1), u(\tau_2)) &\leq d^\beta_\varepsilon(\xi_1, \xi_2) + \beta\left(\|\xi_1 - u(\tau_1)\| + \|\xi_2 - u(\tau_2)\|\right) \\ &\leq \beta\|\xi_2 - \xi_1\| + 2\beta\sqrt{d}\varepsilon. \end{aligned}$$

Consequently we have

$$d_{p,\varepsilon}(\xi_1, \xi_2) = F_{p,\varepsilon}(u) \leq \beta\|\xi_2 - \xi_1\| + 2\beta\sqrt{d}\varepsilon^{1-p} + 2\beta\sqrt{d}\varepsilon.$$

Continuing from (3.65) and applying the triangle inequality with lemma 3.21, we have that,

$$\begin{aligned} F_{p,\varepsilon}(u) \geq \int_{\tau_1}^{\tau_2} a_{p,\varepsilon}\left(\frac{u(\tau)}{\varepsilon}\right)\|u'(\tau)\|\mathrm{d}\tau &= d_{p,\varepsilon}(u(\tau_1), u(\tau_2)) \\ &= d^\beta_\varepsilon(u(\tau_1), u(\tau_2)) \\ &\geq d^\beta_\varepsilon(\xi_1, \xi_2) - d^\beta_\varepsilon(\xi_1, u(\tau_1)) - d^\beta_\varepsilon(\xi_2, u(\tau_2)). \end{aligned}$$

Using (3.50) it follows that

$$F_{p,\varepsilon}(u) \geq \|\xi_2 - \xi_1\| - 2\beta\sqrt{d}\varepsilon.$$

Hence the bounds are illustrated. The remaining case when $T = [0, 1]$ follows in a similar manner. □

The following lemma improves pointwise convergence to local uniform convergence as in the uniformly bounded case. The key here is that we are still close to the uniformly bounded case, due to our improved growth bounds.

Lemma 3.33. *If the metrics $d_{p,\varepsilon}(\xi_1, \xi_2)$ converge pointwise to $\psi(\xi_2 - \xi_1)$ then they converge locally uniformly.*

Proof. We follow the proof of [BPF01, Proposition 2.3]. Take $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$ and let $(x_\varepsilon, y_\varepsilon)$ be a sequence converging to (x, y). Then

$$\begin{aligned} \lim_{\varepsilon\to 0}|d_\varepsilon(x_\varepsilon, y_\varepsilon) - \psi(y - x)| &\leq \lim_{\varepsilon\to 0}|d_\varepsilon(x_\varepsilon, y_\varepsilon) - d_\varepsilon(x, y)| \\ &\quad + \lim_{\varepsilon\to 0}|d_\varepsilon(x, y) - \psi(y - x)| \\ &\leq \lim_{\varepsilon\to 0} C\left(|x_\varepsilon - x| + |y_\varepsilon - y| + 2C\varepsilon^{1-p}\right), \end{aligned}$$

by repeating the proof of lemma 3.11 with the bounds in lemma 3.32. Since the point (x, y) and sequence $\{(x_\varepsilon, y_\varepsilon)\}_{\varepsilon>0}$ are arbitrary, this implies the local uniform convergence required. □

We are now in a position to prove one of our main homogenisation results, using a modification of the method in [BPF01, Theorem 3.1].

Theorem 3.34. *If the induced metrics $d_{p,\varepsilon}$ converge locally uniformly to a norm ψ on $\mathbb{R}^d$, then the sequence of functionals $F_{p,\varepsilon}$ defined on $\mathcal{A}(\xi_1, \xi_2)$ Γ-converge with respect to the $L^\infty(0,1)$ norm topology to*

$$\int_0^1 \psi(u'(\tau))\, d\tau,$$

for all $\xi_1, \xi_2 \in \mathbb{R}^d$.

Proof. Part 1: Lim-inf inequality. Let $(\varepsilon_k)_{k=1}^\infty \subset (0,\infty)$ converge to zero. Fix $u \in \mathcal{A}(\xi_1, \xi_2)$ and let $\mathcal{A}(\xi_1, \xi_2) \ni u_{\varepsilon_k} \to u$ in $L^\infty(0,1)$ as $k \to \infty$, then $u_{\varepsilon_k} \to u$ pointwise as $k \to \infty$. Let $\pi_N = \{\tau_0, ..., \tau_N\}$ be a partition of $[0,1]$ such that $|\tau_j - \tau_{j+1}| = 1/N$ for $j = 1, ..., N$. Then

$$\begin{aligned} F_{\varepsilon_k}(u_{\varepsilon_k}) &= \sum_{i=1}^N \int_{\tau_{i-1}}^{\tau_i} a_{p,\varepsilon}\left(\frac{u_{\varepsilon_k}(\tau)}{\varepsilon_k}\right) \|u'_{\varepsilon_k}(\tau)\| \mathrm{d}\tau \\ &\geq \sum_{i=1}^N d_{p,\varepsilon_k}(u_{\varepsilon_k}(\tau_{i-1}), u_{\varepsilon_k}(\tau_i)), \end{aligned}$$

using the invariance of length under reparameterisations. Therefore applying the triangle inequality for the induced metric, and lemma 3.32,

$$\begin{aligned} F_{\varepsilon_k}(u_{\varepsilon_k}) &\geq \sum_{i=1}^N d_{p,\varepsilon_k}\left(u(\tau_{i-1}), u(\tau_i)\right) - d_{p,\varepsilon_k}\left(u_{\varepsilon_k}(\tau_i), u(\tau_i)\right) - d_{p,\varepsilon_k}\left(u(\tau_{i-1}), u_{\varepsilon_k}(\tau_{i-1})\right) \\ &\geq \sum_{i=1}^N d_{p,\varepsilon_k}\left(u(\tau_{i-1}), u(\tau_i)\right) - \beta\|u_{\varepsilon_k}(\tau_i) - u(\tau_i)\| \\ &\quad - \beta\|u(\tau_{i-1}) - u_{\varepsilon_k}(\tau_{i-1})\| - 2C\varepsilon_k^{1-p}. \end{aligned}$$

Taking the limit as $k \to \infty$ and using theorem 3.30 gives that

$$\liminf_{k\to\infty} F_{\varepsilon_k}(u_{\varepsilon_k}) \geq \sum_{i=1}^N \psi\left(u(\tau_i) - u(\tau_{i-1})\right).$$

Using the $1-$homogeneity of ψ gives

$$\liminf_{k\to\infty} F_{\varepsilon_k}(u_{\varepsilon_k}) \geq \sum_{i=1}^{N} \psi\left(\frac{u(\tau_i)-u(\tau_{i-1})}{|\tau_i-\tau_{i-1}|}\right)|\tau_i-\tau_{i-1}| = \int_0^1 \psi(u_N'(\tau))\mathrm{d}\tau,$$

where u_N is the linear interpolation of u on π_N. Sending $N\to\infty$, and applying the dominated convergence theorem, proves (3.1).

Part 2: Lim-sup inequality. Now to verify (3.2). Fix $u\in\mathcal{A}(\xi_1,\xi_2)$. Then choosing a sequence $(M_k)_{k=1}^{\infty}\subset\mathbb{N}$ such that

$$\lim_{k\to\infty} M_k \sup_{\xi_1,\xi_2\in K} |d_{\varepsilon_k}(\xi_1,\xi_2)-\psi(\xi_2-\xi_1)| = 0,$$

where $K\subset\subset\mathbb{R}^d$ such that $\mathrm{Image}(u)\subset\mathrm{int}(K)$. Let $\pi_{M_k}=\{\tau_0,...,\tau_{M_k}\}$ be a partition of $[0,1]$ such that $|\tau_j-\tau_{j+1}|=1/M_k$ for $j=1,...,M_k$. Define the function u_{ε_k} by

$$u_{\varepsilon_k}(\tau) = u(\tau)+\mathrm{argmin}_{w\in W_0^{1,\infty}(0,1)}\int_{\tau_{i-1}}^{\tau_i} a_{p,\varepsilon_k}\left(\frac{u(\tau)+w(\tau)}{\varepsilon}\right)\|u'(\tau)+w'(\tau)\|\mathrm{d}\tau \quad (3.66)$$

for $\tau\in[\tau_{i-1},\tau_i]$ in the partition π_{M_k}. Clearly, by construction, $u_{\varepsilon_k}\in\mathcal{A}(\xi_1,\xi_2)$. First we prove that $u_{\varepsilon_k}\to u$ in $L^\infty(0,1)$. Fix k and $t\in[0,1]$ and suppose that $\tau\in[\tau_{i-1},\tau_i]$. Then

$$\|u_{\varepsilon_k}(\tau)-u(\tau)\| \leq \|u_{\varepsilon_k}(\tau)-u(\tau_{i-1})\| + \|u(\tau_{i-1})-u(\tau)\|. \quad (3.67)$$

By lemma 3.32 it holds that

$$\begin{aligned}
\alpha\|u_{\varepsilon_k}(\tau)-u(\tau_{i-1})\| &\leq d_{p,\varepsilon_k}(u_{\varepsilon_k}(\tau),u(\tau_{i-1}))+C\varepsilon_k^{1-p},\\
&\leq d_{p,\varepsilon_k}(u_{\varepsilon_k}(\tau_i),u(\tau_{i-1}))+C\varepsilon_k^{1-p}, \text{ since } u_{\varepsilon_k} \text{ is a geodesic},\\
&\leq \beta|u_{\varepsilon_k}(\tau_i)-u_{\varepsilon_k}(\tau_{i-1})|+2C\varepsilon_k^{1-p},\\
&= \beta\|u(\tau_i)-u(\tau_{i-1})\|+2C\varepsilon_k^{1-p}, \text{ by (3.66)}.
\end{aligned}$$

Therefore

$$\|u_{\varepsilon_k}(\tau)-u(\tau)\| \leq \frac{\beta}{\alpha}\|u(\tau_i)-u(\tau_{i-1})\|+\|u(\tau_{i-1})-u(\tau)\|+\frac{2C}{\alpha}\varepsilon_k^{1-p}. \quad (3.68)$$

Fix $\varepsilon>0$. Then, since u is uniformly continuous on $[0,1]$, there exists $\delta>0$ such that $\|u(x)-u(y)\|<\varepsilon$ if $|x-y|<\delta$. Since $|\tau_{i-1}-\tau_i|\to 0$ as $k\to\infty$, for all i, it follows that there exists N such that for $k\geq N$,

$$\|u_{\varepsilon_k}(t)-u(t)\| \leq \left(\frac{\beta}{\alpha}+1\right)\varepsilon_k+\frac{2C}{\alpha}\varepsilon_k^{1-p}. \quad (3.69)$$

Hence $u_{\varepsilon_k} \to u$ in $L^\infty(0,1)$. It remains to show that u_{ε_k} has the desired properties.

$$\begin{aligned}\int_0^1 \psi(u'(\tau))\mathrm{d}\tau &\geq \sum_{i=1}^{M_k} \psi(u(\tau_i) - u(\tau_{i-1})) \\ &\geq \sum_{i=1}^{M_k} d_{p,\varepsilon}(u_{\varepsilon_k}(\tau_i), u_{\varepsilon_k}(\tau_{i-1})) - \\ &\qquad \sum_{i=1}^{M_k} \left|\psi(u(\tau_i) - u(\tau_{i-1})) - d_{p,\varepsilon}(u_{\varepsilon_k}(\tau_i), u_{\varepsilon_k}(\tau_{i-1}))\right|. \end{aligned} \tag{3.70}$$

By construction it holds that

$$\sum_{i=1}^{M_k} d_{p,\varepsilon}(u_{\varepsilon_k}(\tau_i), u_{\varepsilon_k}(\tau_{i-1})) = F_{p,\varepsilon}(u_{\varepsilon_k}),$$

furthermore, for ε sufficiently small,

$$\begin{aligned}&\sum_{i=1}^{M_k} \left|\psi(u(\tau_i) - u(\tau_{i-1})) - d_{p,\varepsilon}(u_{\varepsilon_k}(\tau_i), u_{\varepsilon_k}(\tau_{i-1}))\right| \\ &\qquad\qquad \leq M_k \sup_{\xi_1,\xi_2 \in K} |d_{\varepsilon_k}(\xi_1,\xi_2) - \psi(\xi_2 - \xi_1)|.\end{aligned}$$

Hence by (3.70), the choice of $(M_k)_{k=1}^\infty$ and the liminf inequality,

$$\int_0^1 \psi(u'(\tau))\mathrm{d}\tau \geq \limsup_{k\to\infty} F_{p,\varepsilon}(u_{\varepsilon_k}) \geq \liminf_{k\to\infty} F_{p,\varepsilon}(u_{\varepsilon_k}) \geq \int_0^1 \psi(u'(\tau))\mathrm{d}\tau.$$

Since the choice of sequence $(\varepsilon_k)_{k=1}^\infty$ was arbitrary and the limit is independent of this choice, the Γ-convergence follows. □

It remains to prove the converse, which is significantly easier.

Theorem 3.35. *If the sequence of functionals $F_{p,\varepsilon}$ defined on $\mathcal{A}(\xi_1,\xi_2)$ Γ-converge with respect to the $L^\infty(0,1)$ norm topology to*

$$\int_0^1 \psi(u'(\tau))\,d\tau,$$

for all $\xi_1, \xi_2 \in \mathbb{R}^d$ then the induced metrics converge locally uniformly to a norm on $\mathbb{R}^d$.

Proof. Fix $\xi_1, \xi_2 \in \mathbb{R}^d$ then applying theorem 3.2 it follows that the limit

$$\lim_{\varepsilon\to 0} d_{p,\varepsilon}(\xi_1,\xi_2) = \lim_{\varepsilon\to 0} \inf_{u\in W^{1\infty}(0,1)} \{F_{p,\varepsilon}(u)\colon u(0) = \xi_1, u(1) = \xi_2\}$$

exists, and hence $d_{p,\varepsilon}$ converges pointwise to ψ. By lemma 3.33 the local uniform convergence follows. □

3.4 A New Phenomenon in the Homogenisation of Piecewise Constant Metrics

In this section we evaluate the asymptotic homogenisation formula for a previously unconsidered homogenisation problem.

3.4.1 Problem Set-Up

We compute explicitly the $\Gamma-$limit for the sequence of functionals

$$\int_0^1 a_\rho\left(\frac{u(\tau)}{\varepsilon}\right)\|u'(\tau)\|\mathrm{d}\tau,\ u\in W^{1,1}(0,1), \tag{3.71}$$

where for $\rho\in(\frac{1}{2},1)$ the function a_ρ is defined by

$$a_\rho(x,y):=\begin{cases}\beta, & \text{if } (x,y)\in\Omega_g:=\frac{1}{2}(1-\rho,1+\rho)^2\\ 1, & \text{if } (x,y)\in[0,1]^2\setminus\Omega_g,\end{cases} \tag{3.72}$$

extended periodically to $\mathbb{R}^2$.

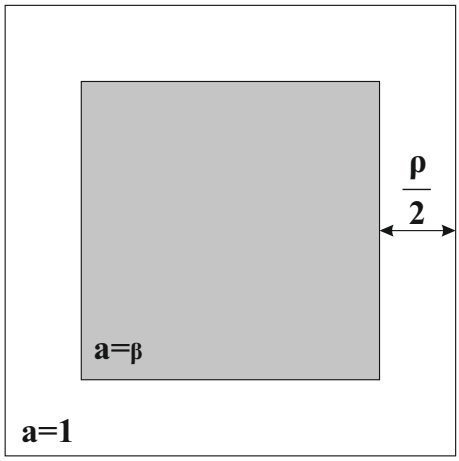

Figure 3-2: *Diagram of the unit cell.*

Recall that the sequence $\Gamma-$converges to a functional of the form

$$\int_0^1 \psi_\rho(u'(\tau))\mathrm{d}\tau, \tag{3.73}$$

where ψ_ρ is defined as

$$\psi_\rho(\xi) = \lim_{\varepsilon\to 0} \min_{u\in\mathcal{A}(0,\xi)} \left\{ \int_0^1 a_\rho\left(\frac{u(\tau)}{\varepsilon}\right) \|u'(\tau)\|\mathrm{d}\tau \right\}. \tag{3.74}$$

The focus of this study is to evaluate (3.74) for (3.71). The case $\rho = 1$ has been previously calculated in [CB03, OTV09], using the Hamilton-Jacobi PDE approach, and the limit $\psi_1(\xi)$ corresponds to the Manhattan norm. Let $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$. The Manhattan norm of ξ is defined as $\|\xi\|_1 := |\xi_1| + |\xi_2|$.

The limit was also calculated in the similar but unbounded case for $\rho = 1$ in [BPF01]. This is to be expected as on the microscopic scale one is confined to moving parallel to the x or y axis when the start end end points are in the region where $a_\rho(x, y) = 1$. It is in fact this property that ensures that geodesics are easy to compute on the microscopic scale. Here we formulate a more general problem, that is, we allow our 'streets' on the microscopic scale to have a non-trivial width, controlled by ρ, see figure 3-3. We can then study the impact of changing this microscopic information on the macroscopic description given by (3.74).

The line of argument for evaluating (3.74) resembles [BD98, Chapter 16], where a checkerboard geometry is considered, with sufficiently high contrast to ensure that one may apply lemma 3.21. The underlying microscopic features of the checkerboard metric make it easy to compute a geodesics by elementary geometric reasoning. In contrast, the problem considered in this paper has a geometry depending on a free parameter and a less restrictive underlying structure; it is thus unclear initially what a geodesic should be, we therefore need additional arguments to determine this. In particular, we reduce the infinite dimensional geodesic problem to a finite dimensional minimisation problem, based on several stages of geometric reasoning. We then solve the minimisation problem.

To the best of our knowledge, no other example gives the homogenised limit as piecewise affine on infinitely many pieces, which is an interesting unobserved phenomenon. Such an example may provide additional insight into the lower contrast checkerboard problem in [ACM09], where the authors experience difficulty in computing the full effective metric for β close to one, but can compute the limit outside of the region where we find infinitely many likes of nondifferentiability accumulating.

This result seems to be the first to include a parameter that modifies the microscopic information, showing explicitly how this effects the macroscopic description given by

(3.74). The effect of varying ρ can be seen in figure 3-8 in subsection 3.4.4. In particular we recover that the limit metric as ρ tends to 1 produces the Manhattan metric as seen in figure 3-7. Additionally, since the limit metric for $\rho \in (\frac{1}{2}, 1)$ produces infinitely many lines of discontinuity, therefore provides reveals a new challenge when trying to determine the limit metric numerically using methods as in [GO04, OTV09].

3.4.2 Computation of the Geodesics

Reduction to shortest path problem on a finite discrete graph

We reduce the computation of a geodesic to that of a shortest path on a discrete graph. In this context a geodesic joining (x_1, y_1) to (x_2, y_2) is a curve u, parameterised on $(0, 1)$, minimising (3.71) subject to $u(0) = (x_1, y_1)$ and $u(1) = (x_2, y_2)$. We compute a specific family of geodesics, for reasons outlined in subsection 3.4.4, using the length functional (3.71). In particular we determine a geodesics joining $\left(\frac{1}{2}(1-\rho), -\frac{1}{2}(1-\rho)\right)$ to $\left(M + \frac{1}{2}(1-\rho), N - \frac{1}{2}(1-\rho)\right)$ for $(M, N) \in \mathbb{N}^2$ with $M > N$. This is clearly equivalent to computing geodesics joining $(0, 0)$ to (M, N) in the shifted length functional

$$\int_0^1 A_\rho(u(\tau))\|u'(\tau)\|\mathrm{d}\tau,\ u \in W^{1,\infty}(0,1), \tag{3.75}$$

where

$$A_\rho(x, y) := a_\rho\left(x + \tfrac{1}{2}(1-\rho), y - \tfrac{1}{2}(1-\rho)\right).$$

For the remainder of this section we consider the latter minimisation problem, for some M, N fixed, as the notation for this problem is less cumbersome. Let us define the sets $TL := (0, 1) + \mathbb{Z}^2$, $TR := (\rho, 1) + \mathbb{Z}^2$, $BL := (0, 1-\rho) + \mathbb{Z}^2$, and $BR := (\rho, 1-\rho) + \mathbb{Z}^2$ corresponding to the top left/right and bottom left/right corners of the squares in Ω_g in the shifted metric, respectively. In addition, recall that we let Ω_g be the set of points (x, y) where $A_\rho(x, y) = \beta$ and $\Omega_\mathrm{w} := \mathbb{R}^2 \setminus \Omega_\mathrm{g}$. See figure 3-3 for an illustration of the notation.

Recall that length functional (3.71) induces a metric on $\mathbb{R}^2$ by setting

$$d_\varepsilon((x_1, y_1), (x_2, y_2)) = \min_{u \in W^{1,\infty}(0,1)} \left\{ \int_0^1 a_\rho\left(\frac{u(\tau)}{\varepsilon}\right) \|u'(\tau)\|\mathrm{d}\tau \colon u(0) = (x_1, y_1), u(1) = (x_2, y_2) \right\}. \tag{3.76}$$

Additionally, the integral in the definition (3.76) may be reparameterised to another

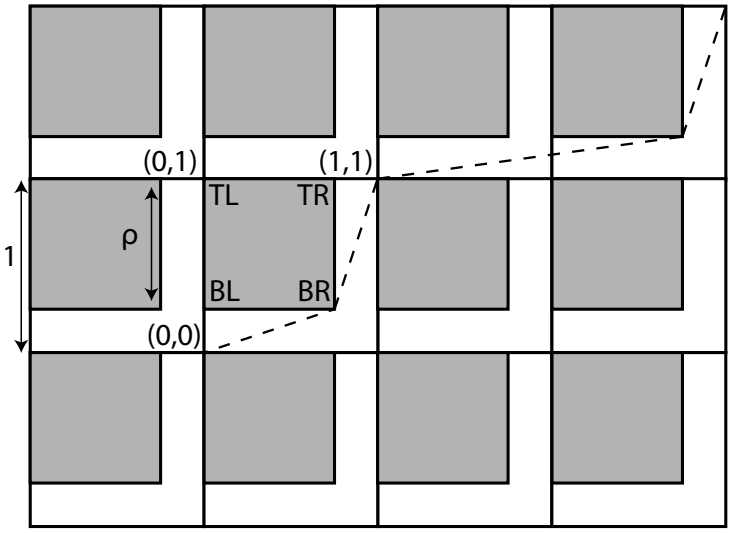

Figure 3-3: *Sketch of the shifted geodesic problem. Elements of the sets TL, TR, BL and BR are indicated. A geodesic for the shifted length functional joining $(0,0)$ to $(3,2)$ is shown. The shaded regions indicate Ω_g.*

interval without changing the value of d by lemma 3.10. Furthermore, d_ε satisfies

$$|(x_1,y_1)-(x_2,y_2)| \leq d_\varepsilon((x_1,y_1),(x_2,y_2)) \leq \beta|(x_1,y_1)-(x_2,y_2)| \tag{3.77}$$

and that the existence of geodesics joining two points is covered by the theorem of Hopf-Rinow [BBI01, Theorem 2.5.28].

The purpose of the next lemma is two-fold. Firstly, a high contrast coefficient exists and $\beta_0 = 2$. Also, it restricts our attention to piecewise affine curves.

Lemma 3.36. *Any geodesic with endpoints in Ω_w does not pass through Ω_g for $\beta > 2$. Furthermore u is piecewise affine.*

Proof. First we prove that the high contrast coefficient is 2. Suppose that there exists a set $S := \Omega_g + (z_1, z_2)$ such that $\text{Image}(u) \cap \varepsilon S \neq \emptyset$. Let G be a connected component of $\text{Image}(u) \cap \varepsilon S$ and let $T := \{\tau \colon u(\tau) \in G\}$. Set

$$u_1 := \inf_{\tau \in T} u(\tau) =: u(\tau_1) \text{ and } u_2 := \sup_{\tau \in T} u(\tau) =: u(\tau_2).$$

Then, define,

$$v(\tau) := \begin{cases} \dfrac{u_\varepsilon(\tau_2) - u_\varepsilon(\tau_1)}{\tau_2 - \tau_1}(\tau - \tau_1) + u_\varepsilon(\tau_1) & \text{if } \tau \in (\tau_1, \tau_2), \\ u_\varepsilon(\tau) & \text{otherwise.} \end{cases}$$

It follows that,

$$\begin{aligned} 2\|u_2 - u_1\| < \beta\|u_2 - u_1\| &= \int_{\tau_1}^{\tau_2} a_\rho(v(\tau))\|v'(\tau)\|\mathrm{d}\tau \\ &= \int_{\tau_1}^{\tau_2} \beta\|v'(\tau)\|\mathrm{d}\tau, \text{ by the convexity of } S, \\ &\leq \int_{\tau_1}^{\tau_2} \beta\|u_\varepsilon'(\tau)\|\mathrm{d}\tau, \text{ by Fermat's principle,} \\ &= \int_{\tau_1}^{\tau_2} a_\rho(u(\tau))\|u_\varepsilon'(\tau)\|\mathrm{d}\tau. \end{aligned} \tag{3.78}$$

Hence we have a lower estimate for the length of the geodesic joining u_1 to u_2. Since S is an open square we consider three cases. Firstly, we can rule out u_1, u_2 are on the same side of S. Since (3.78) shows that we can shorten the curve by not entering S, contradicting the minimality of u_ε. Now suppose that u_1, u_2 are on neighbouring sides of S, let v be the minimum point for,

$$\min\left\{\int_{\tau_1}^{\tau_2} a_\rho(v(\tau))\|v'(\tau)\|\mathrm{d}\tau \colon v(\tau_1) = u_1, v(\tau_2) = u_2, a_\rho(v(\tau)) = 1\right\} = \|u_2 - u_1\|_1,$$

where $\|\cdot\|_1$ is the ℓ^1 norm. By elementary geometric reasoning it holds that,

$$\|u_2 - u_1\|_1 \leq \sqrt{2}\|u_2 - u_1\|,$$

where $\|\cdot\|$ is the Euclidean norm on $\mathbb{R}^d$, and hence by (3.78),

$$\int_{\tau_1}^{\tau_2} a_\rho(v(\tau))\|v'(\tau)\|\mathrm{d}\tau < \int_{\tau_1}^{\tau_2} a_\rho(u(\tau))\|u_\varepsilon'(\tau)\|\mathrm{d}\tau,$$

contradicting the minimality of u_ε. Now suppose that u_1, u_2 are on opposite sides. By elementary geometric reasoning, let v be the minimum point for,

$$\min\left\{\int_{\tau_1}^{\tau_2} a_\rho(v(\tau))\|v'(\tau)\|\mathrm{d}\tau \colon v(\tau_1) = u_1, v(\tau_2) = u_2, a_\rho(v(\tau)) = 1\right\} \leq 2\rho,$$

Furthermore, $\|u_2 - u_1\| \geq \rho$, and therefore by (3.78),

$$\int_{\tau_1}^{\tau_2} a_\rho(v(\tau))\|v'(\tau)\|\mathrm{d}\tau \leq 2\rho \leq 2\|u_2 - u_1\| < \int_{\tau_1}^{\tau_2} a_\rho(u(\tau))\|u_\varepsilon'(\tau)\|\mathrm{d}\tau.$$

Contradicting the minimality of u_ε. The fact that u is piecewise affine follows by identical reasoning to [BD98, Example 16.2]. □

For the remainder of this section it is assumed that $\beta > 2$. Let u be the Lipschitz geodesic joining $(0,0)$ to (M,N) in our metric for $\varepsilon = 1$. Define $I := \{(x,y) \in \mathbb{R}^2 \colon u(T) = (x,y),\ \lim_{\tau\to T^+} u'(\tau) \neq \lim_{\tau\to T^-} u'(\tau)\}$, that is, the points in $\mathbb{R}^2$ where a geodesic changes direction. The next lemma shows that a geodesic only changes direction at the corners of Ω_g.

Lemma 3.37. *It holds that*

$$\big(\mathbb{R}^2 \setminus (TL \cup TR \cup BL \cup BR)\big) \cap I = \emptyset.$$

Proof. Suppose the contrary. By lemma 3.36 any geodesic does not pass through Ω_g, therefore given $x \in I$ it holds that $x \in \mathrm{int}(\Omega_\mathrm{w}) \cup \partial\Omega_\mathrm{w}$. Suppose first that $x \in \mathrm{int}(\Omega_\mathrm{w})$, then there exists an open ball $B_r(x) \subset \mathrm{int}(\Omega_\mathrm{w})$. Let G be the connected component of $\mathrm{Image}(u) \cap B_r(x)$ containing x and let $S := \{\tau \colon u(\tau) \in G\}$. Set $s = \inf S$ and $t = \sup S$ and define

$$v(\tau) := \begin{cases} \dfrac{u(t) - u(s)}{t - s}(\tau - s) + u(s) & \text{if } \tau \in (s,t), \\ u(\tau) & \text{otherwise.} \end{cases}$$

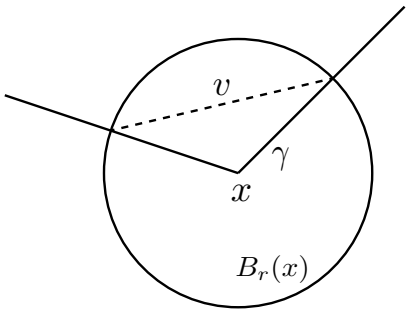

Figure 3-4: *Construction in lemma 3.37*

See figure 3-4 for an illustration of the construction. By construction $v \neq u$ and

$$\int_s^t a_\rho(v(\tau))\|v'(\tau)\|\mathrm{d}\tau < \int_s^t a_\rho(u(\tau))\|u'(\tau)\|\mathrm{d}\tau,$$

contradicting the minimality of u. Now suppose that $x \in \partial\Omega_{\mathrm{w}}$. Since x by assumption is not at a corner of Ω_{g}, there exists a half ball such that the flat edge is contained in $\partial\Omega_{\mathrm{w}}$. Applying the previous argument to the half ball leads in a similar manner to the conclusion that u is not minimal. □

By lemmas 3.36 and 3.37 it follows that a geodesic consists of straight line segments joined at the corners of Ω_{g}. The following lemma reduces the number of potential geodesics to a finite set.

Lemma 3.38. *The image of a geodesic joining* $(0,0)$ *to* (M,N) *is contained in* $[0,M]\times[0,N]$.

Proof. Assume the contrary and suppose further that there exists a point $s \in (0,1)$ such that $u_1(s) < 0$, the other cases are treated similarly. As $u \in C^0(0,1)$ and since $u(1) = (M,N)$, by the intermediate value theorem, there exists $t \in (s,1)$ such that $u_1(t) = 0$, where ℓ_1 denotes the first component of ℓ. Define

$$v(\tau) := \begin{cases} \dfrac{u(t)}{t}\tau & \text{if } \tau \in (0,t), \\ u(\tau) & \text{otherwise.} \end{cases}$$

As in lemma 3.37 it follows that $v \neq u$ and

$$\int_0^t a_\rho(v(\tau))\|v'(\tau)\|\mathrm{d}\tau < \int_0^t a_\rho(u(\tau))\|u'(\tau)\|\mathrm{d}\tau,$$

contradicting the minimality of u. □

The next lemma rules out some corners of Ω_{g} that a geodesic can pass through. More precisely lemma 3.39 shows that a line segment starting at TL must end in a set of BR corners to the right and in the row above.

Lemma 3.39. *Let* $\ell\colon (s,t) \to \mathbb{R}^2$ *be a maximal line segment of a geodesic such that* $\ell(s) = (z_1,z_2) \in TL$ *where* $z_1 \in \{1,\ldots M-1\}$ *and* $z_2 \in \{1,\ldots,N-1\}$. *Then* $\ell(t) = (z_1+Z-(1-\rho), z_2+(1-\rho)) \in BR$ *for* $Z \in \{1,\ldots,M-z_1\}$.

Proof. The proof is split into three cases, depending on the angle at which the line segment leaves TL, denoted by $\theta \in [0,2\pi)$, where $\theta = 0$ is parallel to the x-axis.

Case 1: $\theta \in (\pi/2, 2\pi)$. It is clear that if $\theta \in (3\pi/2, 2\pi)$ then the line segment would continue into Ω_{g}, contradicting lemma 3.36. It remains to rule out that $\theta \in (\pi/2, 3\pi/2]$, which can be achieved using the same construction as in lemma 3.38 to prove there exists a shorter curve.

Case 2: $\theta \in \{0, \pi/2\}$. Suppose that $\theta = \pi/2$; the case $\theta = 0$ follows by a similar argument. In this case, $u(s), u(t) \in \{z_1\} \times [0, N]$. As $u \in C^0(0,1)$ and since $u(0) = (0,0)$, it follows that there exists $r \in (0, s)$ such that $u_1(r) \in \{z_1 - (1-\rho)\} \times [0, N]$. Therefore, applying the same reasoning as in lemma 3.37, we see that a geodesic must consist of straight line segments connecting $u(r)$ to $u(s)$ and $u(s)$ to $u(t)$. However, $u(r), u(s)$ and $u(t)$ form a triangle in the set $[z_1 - (1-\rho), z_1] \times [0, N]$. This contradicts the minimality of u, see figure 3-5.

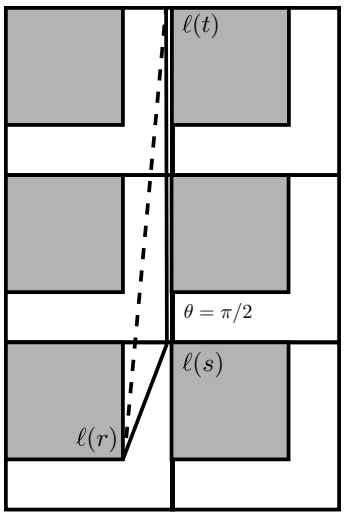

Figure 3-5: *Construction in lemma 3.39 case 2. The vertical solid curve is the line segment ℓ.*

Case 3: $\theta \in (0, \pi/2)$. Suppose first that the line segment connects $\ell(s)$ to any corner not stated in the lemma, consequently $\ell_2(t) - \ell_2(s) \geq 1$, where ℓ_2 is the second component of ℓ. We prove, for $\rho \in (\frac{1}{2}, 1)$, should this line exist, then it necessarily crosses Ω_{g}, contradicting lemma 3.36. Consider the point $u \in (s,t)$ at which $\ell_2(s) + 1 = \ell_2(u)$, which exists by continuity. Then, either $\ell_1(u) \in (P, P+\rho)$ for a $P \in \{0, ..., M-1\}$, in which case by continuity, $\ell(u - \delta) \in \Omega_{\mathrm{g}}$ for δ sufficiently small. Alternatively, $\ell_1(u) \in [P + \rho, P + 1]$ for a $P \in \{0, ...M - 1\}$. Parameterise ℓ over (s, u) as a graph over the x-axis to obtain that $\ell_2(x) = x/\ell_1(u) + \ell_2(s)$ for $x \in (0, \ell_1(u))$. Evaluating ℓ_2 at $x = P + \rho$ gives

$$(1-\rho) + \ell_2(s) < \frac{P+\rho}{P+1} + \ell_2(s) \leq \frac{P+\rho}{\ell_1(u)} + \ell_2(s) \leq 1 + \ell_2(s),$$

if, and only if, $\rho \in (\frac{1}{2}, 1)$. Therefore by continuity, $\ell(u-\delta) \in \Omega_{\mathrm{g}}$ for δ sufficiently small, a contradiction. It remains to rule out that the line segment ends at a BL corner in $W = [\ell_1(s), M] \times (\ell_2(s), \ell_2(s) + (1-\rho)]$. To rule out that the line segment ends in BL, repeat the reasoning of cases 1 and 2 for contradiction. Hence the line segment may only terminate at the BR points of W as stated in the theorem. □

Repeating the reasoning in lemma 3.39 it is possible to show the analogous result for geodesics starting in BR.

Lemma 3.40. *Let $\ell\colon (s,t) \to \mathbb{R}^2$ be a maximal line segment of a geodesic such that $\ell(s) = (z_1 + \rho, z_2 + (1-\rho)) \in BR$ where $z_1 \in \{0, ...M-1\}$ and $z_2 \in \{0, ..., N-1\}$. Then $\ell(t) = (z_1 + 1, z_2 + Z) \in TL$ for $Z \in \{1, ..., N - z_2\}$.*

Lemmas 3.39 and 3.40 state should a geodesic lie in $(0, M) \times (0, N)$ then it necessarily joins points in TL to BR and then BR to TL, in a specific way. We now show that we can extend this property further and rule out that a geodesic lies in $\partial((0, M) \times (0, N))$, except for the end points.

Lemma 3.41. *The image of a geodesic is contained in $(0, M) \times (0, N)$, except for the end points.*

Proof. Reasoning as in the proof of lemma 3.38, it is clear that should a geodesic have a line segment in $\partial((0, M) \times (0, N))$ then it must contain either $(0, 0)$ or (M, N), otherwise it is not minimal. Suppose that the line segment contains $(0, 0)$, the other case is similar. Should the line segment end at $(0, N)$ then by lemma 3.38 it must continue to join $(0, N)$ to (M, N), giving a total length of $M + N$. However, choosing the curve joining $(0, 0)$ to $(M - (1-\rho), \rho) \in BR$ and then onto (M, N) is strictly shorter, therefore the longer curve is not a geodesic. Now suppose that the end of the line segment is $(0, Z) \in \{0\} \times \{1, ..., N-1\}$ (otherwise by previous considerations, the curve is not a geodesic). Then by lemma 3.39 a geodesic must extend as a line segment joining to a point of the form $(Y - (1-\rho), Z + (1-\rho)) \in BR \cap (0, M) \times (0, N)$ for $Y \in \{1, M\}$. Now consider the curve that first joins $(0, 0)$ to $(Y - (1-\rho), 1-\rho) \in BR \cap (0, M) \times (0, N)$, and then continues onto $(Y - (1-\rho), Z + (1-\rho))$, see figure 3-6.

Elementary geometric reasoning shows that the first two line segments of these curves share the same length, and that they both lie in Ω_{w}. However, the latter curve contains a line segment parallel to the y-axis which is forbidden by lemma 3.40 and therefore the curve cannot be minimal. □

We can now identify potential geodesics by a pair of k-tuples. The length of each curve can then be described as a function of those k-tuples. One k-tuple records the

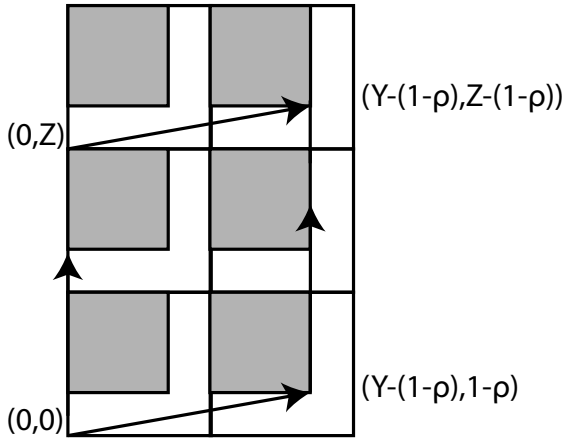

Figure 3-6: *Construction in lemma 3.41.*

distances Z taken by applying lemma 3.39, the other k-tuple records the distances Z from lemma 3.40. Since $(0,0) \in TL$ and $(M,N) \in TL$ and TL connects to BR which connects to TL by lemmas 3.39 and 3.40, it suffices to record such Z to describe the entire structure of the remaining curves.

Lemma 3.42. *The remaining candidate curves can be identified by k-tuples $(m_i)_{i=1}^k$, $(n_i)_{i=1}^k$ with $\sum_{i=1}^k m_i = M$,$\sum_{i=1}^k n_i = N$. The length of a curve with such an identification is*

$$L\left[(m_i)_{i=1}^k,(n_i)_{i=1}^k\right] = \sum_{i=1}^{k} \sqrt{(1-\rho)^2+(m_i-(1-\rho))^2} + \sqrt{(1-\rho)^2+(n_i-(1-\rho))^2}. \tag{3.79}$$

Furthermore, $k \leq N$.

Proof. Starting at $(0,0)$, by lemma 3.41 and reasoning as in lemma 3.39 the candidate geodesic must extend as a line segment joining a point of the form $(Z_1-(1-\rho),1-\rho) \in BR \cap (0,M)\times(0,N)$ for some $Z_1 \in \{1,...,M\}$. This produces a length contribution of $\sqrt{(1-\rho)^2+(m_1-(1-\rho))^2}$, where $m_1 := Z_1$. Since $(M,N) \in TL$, the curve has not yet reached the end point. Therefore, applying lemma 3.40, the candidate geodesic continues as another line segment, connecting to $(m_1,Z_2) \in TL \cap (0,M]\times(0,N]$ for $Z_2 \in \{1,...,N\}$. The contribution to length is $\sqrt{(1-\rho)^2+(n_1-(1-\rho))^2}$, where $n_1 := Z_2$. Now, either $(m_1,n_1) = (M,N)$, in which case we terminate the procedure, or otherwise we may find $m_2 \in \{1,...,M-m_1\}$ and $n_2 \in \{1,...,N-n_1\}$, and so on until $\sum_{i=1}^k m_i = M$, $\sum_{i=1}^k n_i = N$. The procedure obviously ends after $k \leq N$ steps, otherwise we would contradict lemma 3.41. □

The results of this section have demonstrated that a geodesic is reduced to minimising (3.79) over k-tuples in

$$\left\{(m_i)_{i=1}^k, (n_i)_{i=1}^k \in \mathbb{N}^k \colon k \leq N, \sum_{i=1}^k m_i = M, \sum_{i=1}^k n_i = N\right\}. \tag{3.80}$$

Clearly this finite dimensional minimisation problem has a solution.

Minimisation of the length functional

This subsection is dedicated to the calculation of minima for (3.79) over k-tuples in (3.80). For notational convenience set

$$\ell_\rho(x) := \sqrt{(1-\rho)^2 + (x-(1-\rho))^2}. \tag{3.81}$$

To perform this minimisation, we first minimise (3.79) for fixed k and then minimise over k. Lemmas 3.43, 3.44 and 3.45 are technical results to minimise (3.79) for fixed k. Denote by $\lfloor\cdot\rfloor$ the floor function.

Lemma 3.43. *For $x \in [1,\infty)$ and $\rho \in (\frac{1}{2}, 1)$, ℓ_ρ is strictly monotone increasing.*

Proof. A trivial calculus exercise. □

Lemmas 3.44 and 3.45 show that, for fixed k, (3.79) is minimised by distributing the values of the k-tuple equally. Note that the conditions of lemma 3.44 ensure that $|z_1 - z_2| \geq 2$.

Lemma 3.44. *For $z_1, z_2 \in \mathbb{N}$, with $2|(z_1+z_2)$, $z_1 \neq z_2$,*

$$\ell_\rho(z_1) + \ell_\rho(z_2) > 2\,\ell_\rho\left(\frac{z_1+z_2}{2}\right). \tag{3.82}$$

Proof. Suppose without loss of generality that $z_1 \geq (z_1+z_2)/2$ and $z_2 \leq (z_1+z_2)/2$. Observe that by the fundamental theorem of calculus (3.82) holds if and only if

$$\int_{(z_1+z_2)/2}^{z_1} \frac{d\ell_\rho}{dx}(x)\mathrm{d}x - \int_{z_2}^{(z_1+z_2)/2} \frac{d\ell_\rho}{dx}(x)\mathrm{d}x > 0. \tag{3.83}$$

An elementary calculation shows that

$$\frac{d^2\ell_\rho}{dx^2}(x) = \frac{(1-\rho)^2}{\sqrt{(1-\rho)^2 + (x-(1-\rho))^2}^3} > 0, \tag{3.84}$$

for $x \in [1, \infty)$. Thus, by strict monotonicity,

$$\int_{(z_1+z_2)/2}^{z_1} \frac{d\ell_\rho}{dx}(x)\mathrm{d}x > \left(\frac{z_1 - z_2}{2}\right) \frac{d\ell_\rho}{dx}\left(\frac{z_1 + z_2}{2}\right),$$
$$\int_{z_2}^{(z_1+z_2)/2} \frac{d\ell_\rho}{dx}(x)\mathrm{d}x < \left(\frac{z_1 - z_2}{2}\right) \frac{d\ell_\rho}{dx}\left(\frac{z_1 + z_2}{2}\right).$$

Hence, (3.83) and therefore (3.82) holds. □

Lemma 3.45. *For $z_1, z_2 \in \mathbb{N}$, with $2 \nmid (z_1 + z_2)$, $|z_1 - z_2| \geq 2$,*

$$\ell_\rho(z_1) + \ell_\rho(z_2) > \ell_\rho\left(\left\lfloor \frac{z_1 + z_2}{2} \right\rfloor\right) + \ell_\rho\left(\left\lfloor \frac{z_1 + z_2}{2} \right\rfloor + 1\right). \tag{3.85}$$

Proof. Suppose without loss of generality that $z_1 > z_2$. First consider the case when

$$C_1(z_1, z_2) := z_1 - \left\lfloor \frac{z_1 + z_2}{2} \right\rfloor + 1 > 0,$$
$$C_2(z_1, z_2) := \left\lfloor \frac{z_1 + z_2}{2} \right\rfloor - z_2 > 0.$$

Observe that (3.85) holds if, and only if,

$$\int_{\lfloor (z_1+z_2)/2 \rfloor + 1}^{z_1} \frac{d\ell_\rho}{dx}(x)\mathrm{d}x - \int_{z_2}^{\lfloor (z_1+z_2)/2 \rfloor} \frac{d\ell_\rho}{dx}(x)\mathrm{d}x > 0.$$

Then, by strict monotonicity, using (3.84),

$$\int_{\lfloor (z_1+z_2)/2 \rfloor + 1}^{z_1} \frac{d\ell_\rho}{dx}(x)\mathrm{d}x > C_1(z_1, z_2) \frac{d\ell_\rho}{dx}\left(\left\lfloor \frac{z_1 + z_2}{2} \right\rfloor + 1\right),$$
$$\int_{z_2}^{\lfloor (z_1+z_2)/2 \rfloor} \frac{d\ell_\rho}{dx}(x)\mathrm{d}x < C_2(z_1, z_2) \frac{d\ell_\rho}{dx}\left(\left\lfloor \frac{z_1 + z_2}{2} \right\rfloor\right).$$

The claim follows once we have shown that

$$C_1(z_1, z_2) \frac{d\ell_\rho}{dx}\left(\left\lfloor \frac{z_1 + z_2}{2} \right\rfloor + 1\right) - C_2(z_1, z_2) \frac{d\ell_\rho}{dx}\left(\left\lfloor \frac{z_1 + z_2}{2} \right\rfloor\right) > 0. \tag{3.86}$$

By monotonicity, from (3.84), the left hand side of (3.86) is strictly greater than

$$(C_1(z_1, z_2) - C_2(z_1, z_2)) \frac{d\ell_\rho}{dx}\left(\left\lfloor \frac{z_1 + z_2}{2} \right\rfloor\right).$$

Since $C_1(z_1, z_2) - C_2(z_1, z_2) = z_1 + z_2 - 2\lfloor (z_1 + z_2)/2 \rfloor + 1 > 1$ and $\lfloor (z_1 + z_2)/2 \rfloor \geq 1$, it follows that (3.85) holds.

The case $C_1(z_1, z_2) = C_2(z_1, z_2) = 0$ is impossible by our assumption that $|z_1 - z_2| \geq 2$. Since $2 \nmid (z_1 + z_2)$ the cases $C_1(z_1, z_2) = 0, C_2(z_1, z_2) \neq 0$ and $C_2(z_1, z_2) = 0, C_1(z_1, z_2) \neq 0$ also do not arise. □

We now minimise (3.79) over (3.80) subject to $k \leq N$ fixed.

Lemma 3.46. *Let $1 \leq k \leq N$, then we can write $M = \ell_1 k + s_1$, $N = \ell_2 k + s_2$ for $\ell_i, s_i \in \mathbb{N}$ and $0 \leq s_i < k$. Then a pair of k-tuples $(m_i)_{i=1}^k, (n_i)_{i=1}^k \in \mathbb{N}^k$ that minimises the functional*

$$\sum_{i=1}^{k} \ell_\rho(m_i) + \ell_\rho(n_i)$$

subject to

$$\sum_{i=1}^{k} m_i = M, \qquad \sum_{i=1}^{k} n_i = N \tag{3.87}$$

takes the form $m_i = \ell_1$ for $k - s_1$ terms, $m_i = \ell_1 + 1$ for s_1 terms, $n_i = \ell_2$ for $k - s_2$ terms and $n_i = \ell_2 + 1$ for s_2 terms. Furthermore, this solution is unique, up to permutations.

Proof. Suppose, without loss of generality, that the k-tuple $(m_i)_{i=1}^k$ is not of the form $m_i = \ell_1$ for $k - s_1$ terms and $m_i = \ell_1 + 1$ for s_1 terms. Then by constraint (3.87), there exists at least two terms of the k-tuple m_1, m_2 such that $|m_1 - m_2| \geq 2$.

If $2|(m_1 + m_2)$, then by lemma 3.44 it holds that

$$\ell_\rho(m_1) + \ell_\rho(m_2) > 2\,\ell_\rho\left(\frac{m_1 + m_2}{2}\right),$$

contradicting the minimality of the proposed solution. Otherwise $2 \nmid (m_1 + m_2)$, so that by lemma 3.45

$$\ell_\rho(m_1) + \ell_\rho(m_2) > \ell_\rho\left(\left\lfloor \frac{m_1 + m_2}{2} \right\rfloor\right) + \ell_\rho\left(\left\lfloor \frac{m_1 + m_2}{2} \right\rfloor + 1\right),$$

again contradicting the minimality of the proposed solution. The uniqueness up to rearrangement of indices follows from the uniqueness of the representations $M = \ell_1 k + s_1$, $N = \ell_2 k + s_2$. Hence the result holds. □

With a minimiser for each k found, it remains to minimise over k. To achieve this, it suffices to show that increasing k strictly reduces length. Lemmas 3.47 and 3.48 shows replacing the k-tuple with a $k + 1$-tuple leads to a strict reduction in length.

Lemma 3.47. *Let $z_1 \in \mathbb{N}$, suppose $2|z_1$ and $z_1 \geq 2$, then*

$$\ell_\rho(z_1) > 2\,\ell_\rho\left(\frac{z_1}{2}\right). \tag{3.88}$$

Proof. Since $2|z_1$, write $z_1 = 2k$ for some $k \in \mathbb{N}$. Then, (3.88) is equivalent to showing that

$$\int_k^{2k} \frac{d\ell_\rho}{dx}(x)\mathrm{d}x - \ell_\rho(k) > 0. \tag{3.89}$$

By monotonicity, from (3.84), we have that

$$\int_k^{2k} \frac{d\ell_\rho}{dx}(x)\mathrm{d}x - \ell_\rho(k) > k\frac{d\ell_\rho}{dx}(k) - \ell_\rho(k).$$

It is easy to verify that

$$k\frac{d\ell_\rho}{dx}(k) - \ell_\rho(k) = \frac{(1-\rho)(k-2(1-\rho))}{\sqrt{(1-\rho)^2 + (k-(1-\rho))^2}} = \frac{(1-\rho)(k-2(1-\rho))}{\ell_\rho(k)}. \tag{3.90}$$

Furthermore, since $\ell_\rho > 0$, it holds that the right hand side of (3.90) is positive for $k \in \mathbb{N}$. Hence (3.89) holds. □

Lemma 3.48. *Let $z_1 \in \mathbb{N}$, suppose $2 \nmid z_1$ and $z_1 \geq 2$, then*

$$\ell_\rho(z_1) > \ell_\rho\left(\left\lfloor \frac{z_1}{2} \right\rfloor\right) + \ell_\rho\left(\left\lfloor \frac{z_1}{2} \right\rfloor + 1\right). \tag{3.91}$$

Proof. Since $2 \nmid z_1$, write $z_1 = 2k+1$ for some $k \in \mathbb{N}$. Then, (3.91) is equivalent to showing that,

$$\int_{k+1}^{2k+1} \frac{d\ell_\rho}{dx}(x)\mathrm{d}x - \ell_\rho(k) > 0.$$

By monotonicity, from (3.84), we have that

$$\int_{k+1}^{2k+1} \frac{d\ell_\rho}{dx}(x)\mathrm{d}x - \ell_\rho(k) > k\frac{d\ell_\rho}{dx}(k+1) - \ell_\rho(k) > k\frac{d\ell_\rho}{dx}(k) - \ell_\rho(k).$$

Hence continuing from (3.90) in lemma 3.47 completes the proof. □

The following lemma combines lemmas 3.47 and 3.48 to show that the minimal $k+1$-tuples have total length strictly shorter than the minimal k-tuples.

Lemma 3.49. *Let $(z_i)_{i=1}^k$ and $(\tilde{z}_i)_{i=1}^{k+1}$ be a k-tuple and $k+1$-tuple with z_i being a placeholder for either m_i or n_i as in lemma 3.46. Then*

$$\sum_{i=1}^{k} \ell_\rho(z_i) > \sum_{i=1}^{k+1} \ell_\rho(\tilde{z}_i). \tag{3.92}$$

Proof. Suppose that there exists $j \in \{1, ..., k\}$ such that $z_j \geq 2$; without loss of generality assume $j = k$. Define a new $k+1$-tuple by $\hat{z}_i = z_i$ if $i \in \{1, ..., k-1\}$. If $2|z_j$ then set

$\hat{z}_k = \hat{z}_{k+1} = z_j/2$, otherwise set $\hat{z}_k = \lfloor z_j \rfloor /2$ and $\hat{z}_{k+1} = \lfloor z_j \rfloor /2 + 1$. Using lemmas 3.47 or 3.48, it holds that

$$\sum_{i=1}^{k} \ell_\rho(z_i) > \sum_{i=1}^{k+1} \ell_\rho(\hat{z}_i).$$

Furthermore, since $\sum_{i=1}^{k+1} \hat{z}_i = \sum_{i=1}^{k+1} z_i$, by the minimality of $(\tilde{z}_i)_{i=1}^{k+1}$ we have that

$$\sum_{i=1}^{k+1} \ell_\rho(\hat{z}_i) \geq \sum_{i=1}^{k+1} \ell_\rho(\tilde{z}_i).$$

Now consider the case when $z_i \equiv 1$ for all i. This implies that $k = N$, by lemma 3.46, and hence there is no such $k+1$-tuple. □

From lemma 3.49, it is possible to compute $\min L$ explicitly, and the corresponding geodesic curves.

Proposition 3.50. *The length of a geodesic joining* $(0,0)$ *to* (M,N) *is*

$$\mathscr{L}_\rho(M,N) := N\ell_\rho(1) + (M - \lfloor M/N \rfloor N)\,\ell_\rho(\lfloor M/N \rfloor + 1) + (N - M + \lfloor M/N \rfloor N)\,\ell_\rho(\lfloor M/N \rfloor). \quad (3.93)$$

Proof. By lemma 3.49, it is clear that taking $k = N$, with the corresponding N-tuple $(n_i)_{i=1}^N$ where $n_i = 1$ for all i produces curves of minimal length. It follows that the corresponding N-tuple $(m_i)_{i=1}^N$ is also optimal. Writing $M = RN + S$, it holds that $m_i = R$ for $N - S$ terms and $m_i = R + 1$ for S terms. Hence, the minimal length is

$$\mathscr{L}_\rho(M,N) = N\ell_\rho(1) + S\ell_\rho(R+1) + (N-S)\,\ell_\rho(R).$$

Note that $S = M - \lfloor M/N \rfloor N$ and $R = \lfloor M/N \rfloor$, which completes the proof. □

The curve of length (3.93) is not necessarily unique, as the following corollary shows.

Corollary 3.51. *There are precisely* $\binom{N}{M-\lfloor M/N \rfloor N}$ *geodesics joining* $(0,0)$ *to* (M,N).

Proof. The potential source of non-uniqueness stems from the fact that in Proposition 3.50, the N-tuple $(m_i)_{i=1}^N$ is only unique up to a permutation. Hence the result follows. □

The intuition behind this can be seen in figure 3-3. It does not matter whether a geodesic first joins TL to BR over two squares and then the next connection TL to BR is one square, or as can be seen in the figure. This non-uniqueness is reflected in the various permutations of $(m_i)_{i=1}^N$ that we can take.

The next subsection focuses constructing a sequence of geodesics to compute the limit length.

3.4.3 The ε-scaled Problem

The aim of this subsection is to compute a sequence of geodesics, denoted u_ε, for the scaled length functional (3.71). For the ε-dependent problem we choose to compute geodesics joining

$$\left(\varepsilon\tfrac{1}{2}(1-\rho), -\varepsilon\tfrac{1}{2}(1-\rho)\right) \text{ to } \left(M+\varepsilon\tfrac{1}{2}(1-\rho), N-\varepsilon\tfrac{1}{2}(1-\rho)\right) \tag{3.94}$$

for $(M,N) \in \mathbb{N}^2$ with $M > N$. As before, this is equivalent to computing geodesics joining $(0,0)$ to (M,N) in the shifted length functional

$$\int_0^1 A_\rho\left(\frac{u(\tau)}{\varepsilon}\right)\|u'(\tau)\|\mathrm{d}\tau,\ u \in W^{1,1}(0,1), \tag{3.95}$$

where

$$A_\rho(x,y) := a_\rho\left(x-\varepsilon\tfrac{1}{2}(1-\rho), y-\varepsilon\tfrac{1}{2}(1-\rho)\right).$$

For each $\varepsilon > 0$, determining the minimal length of (3.95) is an identical argument to the case when $\varepsilon = 1$ except that all line segments are scaled by a factor ε. Thus for a fixed ε that the length of a geodesic joining $(0,0)$ to $(\varepsilon M, \varepsilon N)$ in (3.95) is $\varepsilon L(M,N)$. Define $L^\varepsilon_\rho(x,y)$ to be the length of a geodesic joining $(0,0)$ to (x,y) in (3.95).

Lemma 3.52. *Let $(x,y) \in \mathbb{Q}^2$, $x > y > 0$, and suppose $x = p/q$, $y = r/s$. Then there exists a sequence $(\varepsilon_k)_{k=1}^\infty$ with $\varepsilon_k \to 0$ as $k \to \infty$ such that*

$$L^{\varepsilon_k}_\rho(x,y) = L_\rho(x,y), \tag{3.96}$$

where $L_\rho(x,y)$ is the extension of (3.93) to $\mathbb{Q}^2$.

Proof. Take $\varepsilon_k = 1/kqs$, $M = kps$ and $N = kqr$. Then by elementary geometric reasoning

$$L^{\varepsilon_k}_\rho(x,y) = \frac{1}{kqs}L_\rho(kps, krq). \tag{3.97}$$

It also holds that $\frac{1}{kqs}L_\rho(kps,krq) = L_\rho(x,y)$ (to show this is a trivial calculation) therefore the result holds. □

3.4.4 The limit metric

In this section we compute the limit metric corresponding to the Γ-limit of the sequence of functionals (3.71).

Lemma 3.53. *Let $(x, y) \in \mathbb{Q}^2$, $x > y > 0$, and suppose $x = p/q$, $y = r/s$. Then the limit metric takes the value*

$$\psi_\rho(x, y) = L_\rho(x, y). \tag{3.98}$$

Proof. By (3.74)

$$\psi_\rho(x, y) = \lim_{i\to\infty} L_\rho^{a,\varepsilon_i}(x, y), \tag{3.99}$$

where

$$L_\rho^{a,\varepsilon}(x, y) = \min_{u\in W^{1,\infty}(0,1)} \left\{ \int_0^1 a_\rho\left(\frac{u(\tau)}{\varepsilon}\right) \|u'(\tau)\| \mathrm{d}\tau \colon u(0) = (0, 0), u(1) = (x, y) \right\}. \tag{3.100}$$

Furthermore, the limit is independent of the choice of $(\varepsilon_i)_{i=1}^\infty$ where $\varepsilon_i \to 0$ as $i \to \infty$ by lemma 3.13. By lemma 3.29 it holds that

$$\begin{aligned}\lim_{\varepsilon\to 0} & d_\varepsilon((0, 0), (x, y)) \\ &= \lim_{\varepsilon\to 0} d_\varepsilon\left((-\varepsilon\tfrac{1}{2}(1-\rho), -\varepsilon\tfrac{1}{2}(1-\rho)), (x - \varepsilon\tfrac{1}{2}(1-\rho), y - \varepsilon\tfrac{1}{2}(1-\rho)\right).\end{aligned}$$

By definition

$$\begin{aligned} d_\varepsilon((0, 0), (x, y)) &= L_\rho^{a,\varepsilon}(x, y), \\ d_\varepsilon\left((-\varepsilon\tfrac{1}{2}(1-\rho), -\varepsilon\tfrac{1}{2}(1-\rho)), (x - \varepsilon\tfrac{1}{2}(1-\rho), y - \varepsilon\tfrac{1}{2}(1-\rho)\right) &= L_\rho^{\varepsilon}(x, y),\end{aligned}$$

Hence $\lim_{\varepsilon\to 0} L_\rho^{a,\varepsilon}(x, y) = \lim_{\varepsilon\to 0} L_\rho^{\varepsilon}(x, y)$. Therefore, by taking $(\varepsilon_k)_{k=1}^\infty$ as in lemma 3.52 it holds that

$$\psi_\rho(x, y) = \lim_{k\to\infty} L_\rho^{a,\varepsilon_k}(x, y) = \lim_{k\to\infty} L_\rho^{\varepsilon_k}(x, y) = L_\rho(x, y),$$

by lemma 3.52. □

It is now possible to construct the limit metric ψ_ρ on $\mathbb{R}^2$.

Theorem 3.54. *The limit metric is given by*

$$\psi_\rho(x, y) = L_\rho(\max\{|x|, |y|\}, \min\{|x|, |y|\}). \tag{3.101}$$

Proof. Use the fact that ψ_ρ is continuous to extend to $(x, y) \in \mathbb{R}^2$, $x \geq y \geq 0$. To extend

to other regions of $\mathbb{R}^2$, follow an identical procedure as before, applying rotations and reflections as necessary. □

Diagrams of the limit metric for different values of ρ are given in figure 3-7. The properties of ψ_ρ are discussed in the next section.

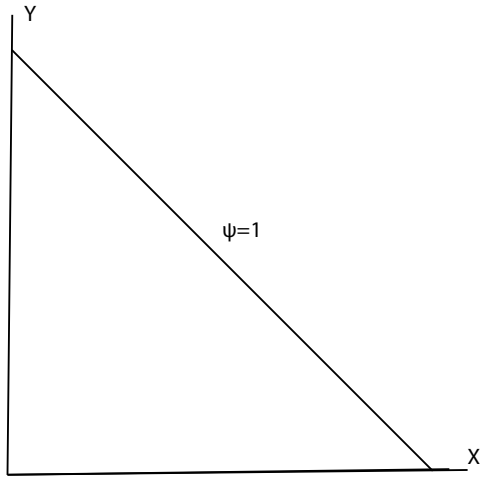

Figure 3-7: *Diagram of the the structure of the set* $\{x \in \mathbb{R}^2 : \psi_1(x) = 1\}$. *The dashed lines are lines of the form* $y = \pm x/k$ *for* $k \in \mathbb{N}$. *The lines of discontinuity accumulate at the* x *and* y *axis. The structure of* ψ_ρ *on other quadrants is obtained by symmetry.*

Properties of the limit metric

It remains to study the structure of ψ_ρ. We show that it is piecewise affine outside of countably many lines of discontinuity.

Lemma 3.55. *The function* ψ_ρ, *restricted to points where* $x > y > 0$, *fails to be differentiable along the lines*

$$y = \frac{x}{k+1},\ k \in \mathbb{N},$$

and $y = x$, $y = 0$. *Furthermore,* ψ_ρ *is piecewise affine.*

Proof. For each (x, y) such that $x > y > 0$ there exists $k \in \mathbb{N}$ such that $1 \leq k \leq x/y < k+1$, and therefore $k = \lfloor x/y \rfloor$, and $x/(k+1) \leq y < x/k$. Consequently, using (3.81), the limit metric takes the form

$$\begin{aligned}\psi_\rho(x, y) &= y\ell_\rho(1) + (x - ky)\,\ell_\rho(k+1) + (y - x + ky)\,\ell_\rho(k) \\ &= \alpha(\rho, k)x + \beta(\rho, k)y,\end{aligned}$$

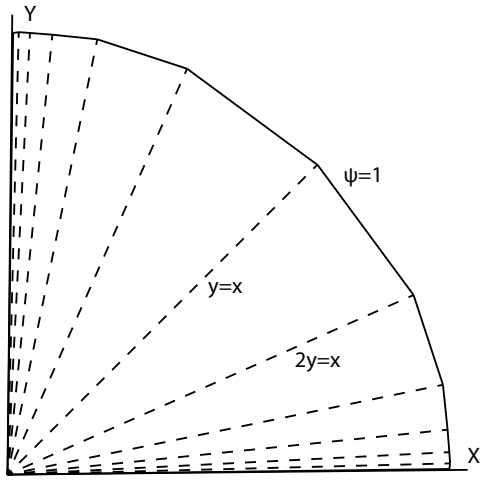

Figure 3-8: *Diagram of the the structure of the set* $\{x \in \mathbb{R}^2 : \psi_\rho(x) = 1\}$. *The dashed lines are lines of the form* $y = \pm x/k$ *for* $k \in \mathbb{N}$. *The lines of discontinuity accumulate at the* x *and* y *axis. The structure of* ψ_ρ *on other quadrants is obtained by symmetry. Here* $\rho \in (\frac{1}{2}, 1)$.

where we have set $\alpha(\rho, k) = \ell_\rho(k+1) - \ell_\rho(k)$ and $\beta(\rho, k) = \ell_\rho(1) + k\left(\ell_\rho(k) - \ell_\rho(k+1)\right) + \ell_\rho(k)$. Clearly, on the set of points such that $x/(k+1) < y < x/k$ it holds that $D\psi(x, y) = (\alpha(\rho, k), \beta(\rho, k)) =: D\psi_k$. This demonstrates that outside of the lines $y = x/(k+1),\ k \in \mathbb{N}$, ψ_ρ is in fact affine. It therefore suffices to verify that the metric is not differentiable along these lines, that is, to show that for $k \in \mathbb{N}$ that $D\psi_k \neq D\psi_{k+1}$, for $k \in \mathbb{N}$. To this end

$$\begin{aligned}
\alpha(\rho, k+1) - \alpha(\rho, k) &= \ell_\rho(k+2) - \ell_\rho(k+1) - \left(\ell_\rho(k+1) - \ell_\rho(k)\right), \\
&= \int_{k+1}^{k+2} \frac{d\ell_\rho}{dx}(x)\mathrm{d}x - \int_k^{k+1} \frac{d\ell_\rho}{dx}(x)\mathrm{d}x \\
&> \frac{d\ell_\rho}{dx}(k+1) - \frac{d\ell_\rho}{dx}(k+1) = 0,
\end{aligned}$$

using the strict monotonicity of $d\ell_\rho/dx$ by (3.84). The lines $y = x$ and $y = 0$ follow with suitable modifications. □

As a consequence of the piecewise affine structure, the following corollary also holds.

Corollary 3.56. *The level sets of* ψ_ρ *are not strictly convex.*

The arguments of this section can be easily adapted to the case where the region of higher length density is on rectangles rather than squares, provided the minimum side length is greater than $1/2$. A similar piecewise affine structure with infinitely many

lines of discontinuity can be derived. The case when $\rho \leq \frac{1}{2}$ would need to be treated via different arguments, since the structure provided by lemmas 3.39 and 3.40 no longer holds. Additionally, the case when $\beta \leq 2$ would require additional reasoning, an example such additional steps for the chessboard geometry can be found in [ACM09].

3.5 Homogenisation of Hamiltonian Dynamics via Metric Methods

The purpose of this section is to formulate an effective description for solutions of

$$\ddot{q}_\varepsilon = -\frac{1}{\varepsilon}\nabla V\left(\frac{q_\varepsilon}{\varepsilon}\right), \tag{3.102}$$

subject to various constraints on q_ε, which will be imposed later, in the limit as $\varepsilon \to 0$. The regularity of V will depend on the problem we study and hence, will also be stated later. As discussed in subsection 3.1.2, the solutions of (3.102) correspond to critical points of the functional

$$\int_0^1 \sqrt{2(E - V(u_\varepsilon(\tau)/\varepsilon))}\|u_\varepsilon'(\tau)\|\mathrm{d}\tau \tag{3.103}$$

when reparameterised by

$$t_\varepsilon(s) = \int_0^s \frac{\|u_\varepsilon'(\tau)\|}{\sqrt{2(E - V(u_\varepsilon(\tau)/\varepsilon))}}\mathrm{d}\tau. \tag{3.104}$$

In this section we denote trajectories of 3.102 by q_ε, which are parameterised by physical time, and denote geodesics of (3.103) by u_ε that are parameterised by arc length. The relationship between u_ε and q_ε being that $u_\varepsilon(s) = q_\varepsilon(t_\varepsilon(s))$. We will apply the homogenisation methods developed in this chapter to the metric formulation (3.103). Consequently we will examine how this can be used to form an effective description of trajectories in (3.102).

The averaging via the Maupertuis principle, using the Jacobi metric (3.103), is trivial in one space dimension. This is since the geodesic connecting two given points is the straight line segment joining these points. All nontrivial information is contained in the time reparameterisation and therefore a simpler approach may be employed to determine the effective behaviour. Specifically, we have an expression for the solution for which the average limit can be calculated explicitly under certain constraints. In addition to the homogenisation of the Hamiltonian fixed energy problem (cf. Theorem 3.58) in one dimension, we also consider the homogenisation of the standard initial value problem.

We observe that when the initial point is fixed, with the additional constraint that the total energy of the system is fixed independently of ε, the solutions q_ε converge uniformly to a single line whose gradient can be described as a function of the total energy. Should one instead fix the initial velocity, rather than the total energy, then depending on how ε approaches 0, it is possible to achieve a range of solutions, all of which are lines. The boundary value problem in one dimension is analysed along with the higher dimensional problems.

In higher dimensions the geodesics have the potential to be more complex, therefore a more general approach is required. The consequence of this is that, while less detailed information is obtained, the range of problems that the general results applies to is increased. We will also demonstrate that other approaches to the homogenisation of Hamiltonian dynamics, as described in the introduction of this chapter, yield the same information as we do here. The key difference being that in the metric formulation has additional structure. Namely that the ε-dependant functionals are 1-homogeneous in the gradient argument, as is the limit functional. It follows that calculating the limit functional numerically can be done on a finite domain, and can easily infer the properties of the limit metric on the whole domain. This should be compared to the alternative of homogenising the action, whereby one will need to perform numerics in the regions of interest in order to obtain accurate information [OTV09].

3.5.1 Homogenisation of Hamiltonian Initial Value Problems in One Space Dimension

Consider the motion of a particle with unit mass travelling in the rapidly oscillating potential $V(q/\varepsilon)$ where $\varepsilon > 0$ is small, and $q \in \mathbb{R}$. The standing assumptions on V are that it is differentiable, V is $P-$periodic and $\max V = 0$. Consider a sequence $\{q_\varepsilon\}_{\varepsilon>0}$ with $q_\varepsilon \colon \mathbb{R} \to \mathbb{R}$ being a solution to (3.102) satisfying

$$\frac{1}{2}\left|\dot{q}_\varepsilon\right|^2 + V\left(\frac{q_\varepsilon}{\varepsilon}\right) = E > 0 \tag{3.105}$$

with initial position $q_\varepsilon(0) = q_a$. We also impose a sign on the initial velocity; without loss of generality we take $\dot{q}_\varepsilon > 0$. The requirement that $\max V = 0$ and $E > 0$ together ensure that the set of points where $E - V(x) > 0$ is $\mathbb{R}$, and hence the Jacobi metric is a Riemannian metric. These requirements are not restrictive.

We define two types of initial value problem one can consider for the homogenisation of Hamiltonian dynamics. The first is the typical initial value problem where the initial position and velocity are fixed, and independent of ε. The second problem is to consider the initial position fixed, however allow the initial velocity to vary with ε, such that the

total energy of the system is fixed and independent of ε.

It is also beneficial to study the boundary value problem, this will be done in a later section.

Convergence of the fixed energy problem

From here it is relatively simple to obtain information about q_ε. We allow the initial velocity to vary with ε is as follows. Fix the total energy $E \in \mathbb{R}$ and assume that our initial velocity is positive. Given the initial position we are then able to solve (3.105) explicitly to obtain

$$\dot{q}_\varepsilon = \sqrt{2\left(E - V\left(\frac{q_\varepsilon}{\varepsilon}\right)\right)}.$$

It is precisely this step that we cannot do for higher dimensional dynamics, and therefore, the subsequent arguments fail to hold for the more general problems. It follows by the separation of variables that

$$t_\varepsilon(q) = \int_{q_a}^{q} \frac{1}{\sqrt{2\left(E - V\left(\frac{\tau}{\varepsilon}\right)\right)}} \mathrm{d}\tau, \tag{3.106}$$

which is well defined as $E > 0$ and $\max V = 0$. Equation (3.106) is also an explicit expression for the inverse of q_ε. To show the uniform convergence of the solutions q_ε for a fixed total energy E, it is easier, given our explicit representation of t_ε, to determine the uniform convergence of t_ε and show that this in turn implies the uniform convergence of the solutions q_ε. For notational convenience define

$$\sigma(E) := \frac{1}{P}\int_0^P \frac{1}{\sqrt{2(E - V(\tau))}} \mathrm{d}\tau,$$

later we will see that this quantity describes the slope of the limit solution as a function of total energy.

Lemma 3.57. *For any $E > 0$, the sequence $\{t_\varepsilon\}_{\varepsilon>0}$ described above satisfies*

$$|t_\varepsilon(q) - t_{\mathrm{hom}}(q)| \leq C_E \varepsilon \tag{3.107}$$

with $t_{\mathrm{hom}}(q) := \sigma(E)(q - q_a)$. C_E is a constant depending only on the periodicity of V and E.

Proof. For given $\varepsilon > 0$, then for each $q \in \mathbb{R}$ we have

$$\begin{aligned} |t_\varepsilon(q) - t_{\hom}(q)| &= \left| \int_{q_a}^{q} \left(\frac{1}{\sqrt{2\left(E - V\left(\frac{\tau}{\varepsilon}\right)\right)}} - \sigma(E) \right) \mathrm{d}\tau \right| \\ &= \varepsilon \left| \int_{q_a/\varepsilon}^{q/\varepsilon} \left(\frac{1}{\sqrt{2(E - V(z))}} - \sigma(E) \right) \mathrm{d}z \right|. \end{aligned}$$

If ε is such that $q - q_a > P\varepsilon$, then $\varepsilon^{-1}[q_a, q]$ is the union of a finite number of fundamental periods of V and a remainder set Q_ε whose measure is

$$\mathrm{meas}(Q_\varepsilon) = \frac{q - q_a}{P\varepsilon} - \left[\left(\frac{q - q_a}{P\varepsilon} \right) - 1 \right] \leq 2.$$

Hence,

$$\begin{aligned} \left| \varepsilon \int_{q_a/\varepsilon}^{q/\varepsilon} \left(\frac{1}{\sqrt{2(E - V(z))}} - \sigma(E) \right) \mathrm{d}z \right| &= \left| \varepsilon \int_{Q_\varepsilon} \left(\frac{1}{\sqrt{2(E - V(z))}} - \sigma(E) \right) \mathrm{d}z \right| \\ &\leq \varepsilon \int_{Q_\varepsilon} \left| \frac{1}{\sqrt{2(E - V(z))}} - \sigma(E) \right| \mathrm{d}z \\ &\leq \frac{4\varepsilon}{\sqrt{2E}}, \end{aligned}$$

using the fact that $V \leq 0$ and the definition of $\sigma(E)$. Otherwise $q - q_a \leq P\varepsilon$ and therefore

$$\left| \int_{q_a}^{q} \left(\frac{1}{\sqrt{2\left(E - V\left(\frac{\tau}{\varepsilon}\right)\right)}} - \sigma(E) \right) \mathrm{d}\tau \right| \leq \frac{2P\varepsilon}{\sqrt{2E}}.$$

Therefore the claim follows, and the sequence $\{t_\varepsilon\}_{\varepsilon>0}$ converges uniformly to $t_{\hom}(q)$. □

From Lemma 3.57 it follows that the initial value problem, where the initial velocity depends on the fixed energy E, is well posed, and that the solutions converge to a unique limit.

Theorem 3.58. *Consider the Hamiltonian initial value problem for the motion of a particle with unit mass,*

$$\begin{cases} \ddot{q}_\varepsilon(t) = -\dfrac{1}{\varepsilon} V'\left(\dfrac{q_\varepsilon(t)}{\varepsilon}\right), \\ q(0) = q_a,\ \dot{q}(0) = p_\varepsilon, \end{cases} \tag{3.108}$$

where $p_\varepsilon > 0$ is chosen to satisfy (3.105) *for a fixed $E > 0$. Then the sequence $\{q_\varepsilon\}_{\varepsilon>0}$*

converges uniformly on $\mathbb{R}$*, with*

$$\sup_{t\in\mathbb{R}}\left|q_\varepsilon(t)-\frac{t}{\sigma(E)}-q_a\right|\leq\frac{C_E}{\sigma(E)}\varepsilon,$$

where C_E *is as in Lemma 3.57.*

Proof. Fix $t\in\mathbb{R}$. Then

$$\begin{aligned}\left|q_\varepsilon(t)-\frac{t}{\sigma(E)}-q_a\right|&=\left|q_\varepsilon(t)-\frac{t_\varepsilon(q_\varepsilon(t))}{\sigma(E)}-q_a\right|\\&=\frac{1}{\sigma(E)}\left|\sigma(E)(q_\varepsilon(t)-q_a)-t_\varepsilon(q_\varepsilon(t))\right|\\&\leq\frac{C_E}{\sigma(E)}\varepsilon,\end{aligned}$$

the last step following from (3.107). Taking the supremum over $t\in\mathbb{R}$ proves the claim. □

In the one dimensional setting it is possible to determine further, exact, information about the limit solutions in Theorem 3.58. For instance, the following lemma shows that the slopes of q_ε tend to 0 as E approaches the critical value 0.

Lemma 3.59. *It holds that* $\sigma(E)\to\infty$ *as* $E\to 0$*.*

Proof. For any $E>0$, since $-V\geq 0$,

$$\begin{aligned}\sigma(E)=\frac{1}{P}\int_0^P\frac{1}{\sqrt{2(E-V(\tau))}}\mathrm{d}\tau&\geq\frac{1}{P}\int_0^P\frac{1}{\sqrt{2E}}\mathrm{d}\tau\\&=\frac{1}{\sqrt{2E}}.\end{aligned}$$

It follows now taking $E\to 0$ that $\sigma(E)\to\infty$. □

Before proceeding to show that the standard initial value problem is ill-posed, we offer a few remarks. Lemma 3.59 also suggests a rate for the divergence of σ. It is a sharp rate, as we can take $V=0$. Recall only the cases where $E>0$ were considered. Should $E\leq 0$ then the corresponding solutions of (3.102) are either separatrices or periodic trajectories with amplitudes of order ε, therefore in the limit $\varepsilon\to 0$ solutions necessarily converge to constant solutions (and hence zero slope). As a consequence one can see from Lemma 3.59 and Theorem 3.58 that the slope of the limiting solution depends continuously on E.

Convergence of the standard initial value problem

Theorem 3.58 can be applied to the standard initial value problem to show a type of non-convergence result. It is clear that the energy E, which is fixed, appears as part of the limiting solution. Since in this situation there is a bijective correspondence between the initial momentum and ε it means that this limit is uniquely characterised by the given data. On the other hand, by fixing the initial position and momentum but not the energy E allows for the value of E to depend on ε. In particular, one can select those ε that give rise to the same E, which exist due to the periodicity of the potential V, and consequently it is possible to apply the previous results to conclude convergence to the line whose slope corresponds to the given E. Therefore it can be seen that different choices of subsequences give rise to different limits. In one dimension there exists a range of possible limits which are characterised in the following theorem.

Theorem 3.60. *Consider the Hamiltonian initial value problem for the motion of a particle with unit mass,*

$$\begin{cases} \ddot{q}_\varepsilon(t) = -\frac{1}{\varepsilon} V'\left(\frac{q_\varepsilon(t)}{\varepsilon}\right), \\ q(0) = q_a \neq 0,\ \dot{q}(0) = p_a > 0. \end{cases} \tag{3.109}$$

Then there exists an interval I of positive length, depending only on $\min V$ and p_a, such that for all $S \in I$ there exists a sequence $\{\varepsilon_k\}_{k\in\mathbb{N}}$ converging to zero such that the corresponding solutions q_{ε_k} converge uniformly to the line $q(t) = St + q_a$. In particular if $p_a < \sqrt{2\max(-V)}$ then $\inf I = 0$.

Proof. Let $M > 0$ be the sharp bound such that $-M \leq V \leq 0$ and suppose that we choose $E > 0$ such that $-M \leq E - \frac{1}{2}p_a^2 \leq 0$. Then let us solve the energy identity

$$\frac{1}{2}|p_a|^2 + V\left(\frac{q_a}{\varepsilon}\right) = E \tag{3.110}$$

for ε to obtain

$$\varepsilon_k = \frac{q_a}{V_0^{-1} + Pk}, \quad k \in \mathbb{N}$$

where V_0^{-1} satisfies $V(V_0^{-1}) = E - \frac{1}{2}p_a^2$, which exists by our choice of E. Then fixing any $E \in (\min\{(\frac{1}{2}p_a^2 - M), 0\}, \frac{1}{2}p_a^2]$ determines a sequence $\{\varepsilon_k\}_{k\in\mathbb{N}}$, depending on E. The sequence $\{\varepsilon_k\}_{k\in\mathbb{N}}$ clearly converges to 0. Then each solution of the IVP (3.109) has energy $E > 0$. One can apply the result from Theorem 3.58 to conclude that q_{ε_k} converges to

$$q(t) = \frac{t}{\sigma(E)} + q_a.$$

This can be repeated for any $E \in (\min\{(\frac{1}{2}p_a^2 - M), 0\}, \frac{1}{2}p_a^2]$. The set I can thus be defined as

$$I := \left\{ \frac{1}{\sigma(E)} \;\middle|\; E \in (\min\{(\tfrac{1}{2}p_a^2 - M), 0\}, \tfrac{1}{2}p_a^2] \right\}.$$

If $p_a < \sqrt{2\max(-V)}$, then $\min\{(\frac{1}{2}p_a^2 - M), 0\} = 0$ and hence to prove $\inf I = 0$ we need to show

$$\lim_{E \to 0} \frac{1}{\sigma(E)} = 0,$$

which holds by Lemma 3.59. □

We note that in the above it is required that $q_a \neq 0$ to ensure that the energy identity (3.110) can be solved for ε. The above theorems and proofs hold, *mutatis mutandis*, for the case where negative initial velocities instead of positive ones are selected.

Comparison with Hamilton-Jacobi theory

Theorem 3.58 gives a complete description of the solution trajectories for (3.102) in the limit $\varepsilon \to 0$. We now show that the averaging approach via Hamilton-Jacobi theory [LPV88] recovers the effective limit. The general approach to averaging via Hamilton-Jacobi theory is summarised in [Eva92].

For the Hamiltonian

$$H(p, q) = \frac{1}{2}\|p\|^2 + V(q) \tag{3.111}$$

in one dimension, it has been observed that a precise form for the effective Hamiltonian can be derived [LPV88, E91, Gom00, GO04]

$$\mathscr{H}^{\mathrm{hom}}(p) := \begin{cases} 0 & \text{if } |p| \leq \dfrac{1}{P}\displaystyle\int_0^P \sqrt{-2V(s)}\,\mathrm{d}s, \\ \alpha & \text{if } |p| = \dfrac{1}{P}\displaystyle\int_0^P \sqrt{2(\alpha - V(s))}\,\mathrm{d}s. \end{cases} \tag{3.112}$$

Figure 3-9 is a sketch for the case $V(x) = -\sin(2\pi x)^2$. The flat spot occurring in the in the Hamiltonian, credited to the lack of differentiability in the corresponding Lagrangian [BD78], has been described as 'trapping' [E91], since the form of Hamilton's ODEs take for $\mathscr{H}^{\mathrm{hom}} = 0$ the form

$$\dot{q} = 0 \text{ and } \dot{p} = 0.$$

In the context of the one dimensional Maupertuis principle the reason for trapping is clear; the particle becomes 'trapped' as the potential wells are drawn arbitrarily close together.

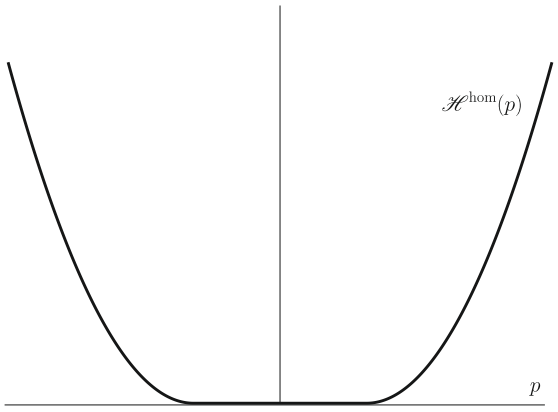

Figure 3-9: *A sketch of the effective Hamiltonian for* $V(x) = -\sin(2\pi x)^2$ *according to [LPV88].*

The solution of the Hamilton-Jacobi PDE

$$u_t = \mathscr{H}^{\hom}(u_q), \tag{3.113}$$

if obtained via separation of variables. Since the Hamiltonian is independent of time, (3.113) reduces to

$$\mathscr{H}^{\hom}(u_q) = E \in \mathbb{R},$$

where E is energy. It is clear that from (3.112) that for any $E > 0$ one has

$$u_q(q) = \pm\frac{1}{P}\int_0^P \sqrt{2(E - V(s))}\,\mathrm{d}s.$$

We recall Jacobi's theorem, paraphrased for this context, from [Eva98, Section 3.2, Theorem 1] that connects the function u to trajectories of (3.102).

Theorem 3.61 (Jacobi)**.** *Let* $u \in C^2(\mathbb{R})$ *solve* (3.113)*. Assume that* q *satisfies the ODE*

$$\dot{q} = \frac{d}{dp}\mathscr{H}^{\hom}(p), \tag{3.114}$$

where $p = u_q \circ q$*. One has that* q *and* p *are solutions to the characteristic equations for* (3.113)*, namely the Hamilton's ODEs.*

With this in mind the final theorem for the one dimensional problem can be stated.

Theorem 3.62. *The effective trajectories of* (3.102) *in one dimension derived through the Hamilton-Jacobi approach coincide with the trajectories obtained in Theorem 3.58.*

Proof. Since u_q is constant we immediately obtain the form of p that solves (3.114). The inverse function theorem, assuming that the momentum is larger than

$$\frac{1}{P}\int_0^P \sqrt{-2V(s)}\,\mathrm{d}s,$$

gives, via the explicit formula for $\mathscr{H}^{\mathrm{hom}}$, that

$$\frac{\mathrm{d}}{\mathrm{d}p}\mathscr{H}^{\mathrm{hom}}(p) = \frac{1}{p'\left(\mathscr{H}^{\mathrm{hom}}(p)\right)}, \tag{3.115}$$

where we have defined

$$p(\alpha) := \frac{1}{P}\int_0^P \sqrt{2(\alpha - V(s))}\,\mathrm{d}s, \text{ for } \alpha > 0.$$

Trivially $\mathscr{H}^{\mathrm{hom}}(u_q) = \mathscr{H}^{\mathrm{hom}}(p) = E$ and therefore

$$\dot{q} = \frac{1}{p'\left(\mathscr{H}^{\mathrm{hom}}(p)\right)} = \frac{1}{p'\left(E\right)},$$

By integration, incorporating the initial position, gives

$$q(t) = \frac{t}{p'\left(E\right)} + q_a.$$

The fact that $\sigma(E) = p'\left(E\right)$ follows from differentiating p as defined above. Hence the result follows. □

We have just shown that

$$\frac{\mathrm{d}}{\mathrm{d}p}\mathscr{H}^{\mathrm{hom}}(p) = \frac{1}{\sigma(E)} = \frac{1}{\sigma(\mathscr{H}^{\mathrm{hom}}(p))}.$$

Since $\frac{\mathrm{d}}{\mathrm{d}p}\mathscr{H}^{\mathrm{hom}}(p) > 0$ for $p > \frac{1}{P}\int_0^P \sqrt{-2V(s)}\,\mathrm{d}s$ it follows that $\mathscr{H}^{\mathrm{hom}}(p) \to 0$ as $p \searrow \frac{1}{P}\int_0^P \sqrt{-2V(s)}\,\mathrm{d}s$. Hence by Lemma 3.59

$$\frac{\mathrm{d}}{\mathrm{d}p}\mathscr{H}^{\mathrm{hom}}(p) \to 0 \text{ as } p \searrow \frac{1}{P}\int_0^P \sqrt{-2V(s)}\,\mathrm{d}s.$$

Consequently the effective Hamiltonian (3.112) is continuously differentiable.

3.5.2 Homogenisation of Hamiltonian Boundary Value Problems

Consider the motion of a particle with unit mass travelling in the rapidly oscillating potential $V(q/\varepsilon)$ where $\varepsilon > 0$ is small, and $x \in \mathbb{R}^d$, $d \geq 1$. The standing assumptions on V are that it is differentiable, V is $[0,1]^d-$periodic and $\max V = 0$. Consider a sequence $\{q_\varepsilon\}_{\varepsilon>0}$ with $q_\varepsilon \colon \mathbb{R} \to \mathbb{R}$ being a solution to (3.102) satisfying

$$\frac{1}{2}|\dot{q}_\varepsilon|^2 + V\left(\frac{q_\varepsilon}{\varepsilon}\right) = E > 0 \tag{3.116}$$

with boundary values $q_\varepsilon(0) = q_a$ and $q_\varepsilon(T_\varepsilon) = q_b$. The aim of this section is to describe and justify a procedure to determine an effective description of q_ε for small ε via the Maupertuis principle. Since $E > \max V$ it follows that (3.103) describes the length functional for a Riemannian manifold, and hence, applying theorem 3.17 we have the following result.

Theorem 3.63 (Homogenisation of the BVP for the Jacobi Metric)**.** *Let $q_a, q_b \in \mathbb{R}^d$. The functionals*

$$F_\varepsilon(u) = \int_0^1 \sqrt{2(E - V(u(\tau)/\varepsilon))}\|u'(\tau)\|\,d\tau$$

Γ-converge, on $\mathcal{A}(\xi_1, \xi_2)$ with respect to the $L^\infty(0,1)$ topology, to a functional of the form

$$F(u) := \int_0^1 \psi(u'(\tau))\,d\tau,$$

where ψ, as defined in lemma 3.13, defines a norm on $\mathbb{R}^d$. Furthermore, a function u minimises F, if and only if

$$\int_0^1 \psi(u'(\tau))\,d\tau = \psi\left(\int_0^1 u'(\tau)\,d\tau\right).$$

The functional F always has a minimiser in $\mathcal{A}(\xi_1, \xi_2)$ which is $\bar{u}(\tau) = (\xi_2 - \xi_1)\tau + \xi_1$.

The first part of the procedure to study the convergence of the geodesics u_ε joining q_a to q_b as $\varepsilon \to 0$.

Lemma 3.64. *Let $q_a, q_b \in \mathbb{R}^d$. Then the minimisers of*

$$F_\varepsilon(u) = \int_0^1 \sqrt{2(E - V(u(\tau)/\varepsilon))}\|u'(\tau)\|\,d\tau$$

such that $u(0) = q_a$ and $u(1) = q_b$, when parameterised proportionally to arc length, form a precompact set in $L^\infty(0,1)$.

Proof. It is well known that minimisers of F_ε, when parameterised by arc length, are also minimisers of

$$G_\varepsilon(u) := \int_0^1 2(E - V(u(\tau)/\varepsilon))\|u'(\tau)\|^2 \mathrm{d}\tau.$$

The following trivial estimate holds

$$\int_0^1 2E\|u'(\tau)\|^2 \mathrm{d}\tau \leq G_\varepsilon(u) \leq \int_0^1 2\left(E + \max_{[0,1]^d}\{-V(x)\}\right)\|u'(\tau)\|^2 \mathrm{d}\tau. \tag{3.117}$$

Hence taking the infimum of (3.117) over all curves joining q_a to q_b, and observing that the infimum is attained in each case, by the Hopf-Rinow theorem for G_ε [Jos05, Chapter 1], it follows that

$$2E\|q_b - q_a\|^2 \leq G_\varepsilon(u_\varepsilon) \leq 2\left(E + \max_{[0,1]^d}\{-V(x)\}\right)\|q_b - q_a\|^2. \tag{3.118}$$

Consequently $\|u_\varepsilon'\|_{L^2(0,1)}$ is bounded. By the Poincaré inequality, since the u_ε satisfy the same boundary conditions, it follows that $\|u_\varepsilon\|_{W^{1,2}(0,1)}$ is bounded. Applying the Rellich-Kondrachov theorem gives that $\{u_\varepsilon\}_{\varepsilon>0}$ is precompact in $L^\infty(0,1)$. In fact, we can strengthen this result to show that $\|u_\varepsilon\|_{W^{1,\infty}(0,1)}$ is bounded. Since u_ε minimises G_ε, and is parameterised by arc length it follows that $2(E - V(u_\varepsilon(\tau)/\varepsilon))\|u_\varepsilon'(\tau)\|^2$ is constant [Jos05] and $2(E - V(u_\varepsilon(\tau)/\varepsilon))\|u_\varepsilon'(\tau)\|^2 = G_\varepsilon(u_\varepsilon)$, almost everywhere. By (3.118) it follows that

$$2(E - V(u_\varepsilon(\tau)/\varepsilon))\|u_\varepsilon'(\tau)\|^2 \leq 2\left(E + \max_{[0,1]^d}\{-V(x)\}\right)\|q_b - q_a\|^2$$

hence $\|u_\varepsilon'\|_{L^\infty(0,1)} \leq C$, for some C independent of ε. Applying the Poincaré inequality and the Rellich-Kondrachov theorem as before gives that $\|u_\varepsilon\|_{W^{1,\infty}(0,1)}$ is bounded. □

Using lemma 3.64 in combination with theorem 3.63 and 3.2 gives the following result immediately.

Theorem 3.65. *The limit of every convergent subsequence of the sequence of geodesics $\{u_\varepsilon\}_{\varepsilon>0}$ joining q_a to q_b is a minimiser of F; where F is defined in theorem 3.63.*

Before we can relate the limit points of $\{u_\varepsilon\}_{\varepsilon>0}$ to those of the sequence q_ε we need to understand how the re-parameterisations t_ε behave in the limit.

Lemma 3.66. *For $j \in \mathbb{N}$, let $t_{\varepsilon_j} : [0,1] \to [0,\infty)$ be a sequence of reparameterisations corresponding to a convergence sequence of geodesics joining q_a to q_b. Then $\{t_{\varepsilon_j}\}_{j\in\mathbb{N}}$ is precompact in $L^\infty(0,1)$.*

Proof. Recall that

$$t_{\varepsilon_j}(s) = \int_0^s \frac{\|u'_\varepsilon(\tau)\|}{\sqrt{2(E - V(u_{\varepsilon_j}(\tau)/\varepsilon))}} \mathrm{d}\tau.$$

By the Lebesgue differentiation theorem it follows that

$$t'_{\varepsilon_j}(s) = \frac{\|u'_{\varepsilon_j}(s)\|}{\sqrt{2(E - V(u_{\varepsilon_j}(s)/\varepsilon))}},$$

almost everywhere. It was demonstrated in lemma 3.64 that there exists a C independent of ε such that $\|u'_\varepsilon\|_{L^\infty(0,1)} \leq C$. Since $E > \max V$ and V is bounded, by continuity and periodicity, it follows that $\|t'_\varepsilon\|_{L^\infty(0,1)} \leq D$ for some $D \in \mathbb{R}$ independent of ε. To bound $\|t_{\varepsilon_j}\|_{\varepsilon>0}$ observe that

$$\int_0^s \frac{\|u'_{\varepsilon_j}(\tau)\|}{\sqrt{2(E - V(u_{\varepsilon_j}(\tau)/\varepsilon))}} \mathrm{d}\tau \leq s\|t'_\varepsilon\|_{L^\infty(0,1)} \leq D.$$

Applying the Rellich-Kondrachov theorem, using the fact that $\|t_{\varepsilon_j}\|_{W^{1,\infty}(0,1)} \leq 2D$, it follows that $\{t_{\varepsilon_j}\}_{\varepsilon>0}$ is precompact in $L^\infty(0,1)$. □

Recall that the reparameterisation t_{ε_j} will have an image that is also ε-dependent, corresponding to the arrival time of the particle at q_b. The function t_{ε_j} is strictly monotone, as $\sqrt{2(E - V(u_{\varepsilon_j}(s)/\varepsilon))} > 0$ for all s and $\|u'_{\varepsilon_j}(s)\| > 0$ for all s. The latter inequality following from the fact that u_{ε_j} is an arc length parameterised geodesic [BBI01]. Consequently t_{ε_j} is invertible. Since $q_{\varepsilon_j} = u_{\varepsilon_j} \circ t_{\varepsilon_j}^{-1}$, in order to determine the uniform convergence of q_{ε_j}, we must establish the uniform convergence of $t_{\varepsilon_j}^{-1}$. This is so we can make use of the fact that the composition of uniformly convergence sequences of functions is uniformly convergent. Observe that the map $t_{\varepsilon_j}^{-1}$ will have a domain of the form $[0, T_{\varepsilon_j}]$ where T_{ε_j} is the ε-dependant arrival time defined by

$$T_{\varepsilon_j} = \int_0^1 \frac{\|u'_{\varepsilon_j}(\tau)\|}{\sqrt{2(E - V(u_{\varepsilon_j}(\tau)/\varepsilon))}} \mathrm{d}\tau$$

To handle the convergence of functions with varying domains we will use the following result from [BDF91], restated in the context of the functions we study here.

Theorem 3.67. *Let $\{t_{\varepsilon_j}\}_{j=1}^\infty$ be a sequence of injective functions defined on*

$$(\alpha, \beta) \subset \cap_{j=1}^\infty dom(t_{\varepsilon_j}).$$

If $\{t_{\varepsilon_j}\}_{j=1}^{\infty}$ converges uniformly to t_0 on (α,β), and t_0 is a continuous injective function, then $\{t_{\varepsilon_j}^{-1}\}_{j=1}^{\infty}$ converges uniformly to t_0^{-1} on each $(A,B) \subset \cap_{j=1}^{\infty} t_{\varepsilon_j}((\alpha,\beta))$.

Proof. Cf. [BDF91, Theorem 1]. □

In order to apply theorem 3.67 we have to show that any limit point of the sequence is continuous and injective. Continuity is clear, each t_{ε_j} is Lipschitz continuous as shown in the proof of lemma 3.66. We prove injectivity in the next lemma.

Lemma 3.68. *Let t_0 be a limit point of the sequence $\{t_{\varepsilon_j}\}_{j=1}^{\infty}$ with respect to uniform convergence, then it is strictly monotone and therefore injective.*

Proof. Fix $\tau_1, \tau_2 \in (0,1)$, $\tau_2 > \tau_1$ and consider

$$t_{\varepsilon_j}(\tau_2) - t_{\varepsilon_j}(\tau_1) = \int_{\tau_1}^{\tau_2} t'_{\varepsilon_j}(s)\mathrm{d}s = \int_{\tau_1}^{\tau_2} |t'_{\varepsilon_j}(s)|\mathrm{d}s \geq (\tau_2 - \tau_1) \inf_{s\in[0,1]} \|t'_{\varepsilon_j}(s)\|. \tag{3.119}$$

We aim to bound $\|t'_{\varepsilon_j}\|_{L^\infty(0,1)}$ from below by a positive constant. Since $\max V = 0$ it holds that

$$\frac{1}{\sqrt{2(E - V(u_{\varepsilon_j}(s)/\varepsilon))}} \geq \frac{1}{\sqrt{2E}}. \tag{3.120}$$

Furthermore proceeding as in the proof of lemma 3.64, we see that

$$2E\|q_b - q_a\|^2 \leq G_\varepsilon(u_{\varepsilon_j}) \leq 2\left(E + \max_{[0,1]^d}\{-V(x)\}\right)\|q_b - q_a\|^2. \tag{3.121}$$

As u_ε is a sequence of geodesics parameterised by arc length it follows that $G_{\varepsilon_j}(u_{\varepsilon_j}) = 2(E - V(u_{\varepsilon_j}(s)/\varepsilon_j))\|u'_{\varepsilon_j}(s)\|^2 \in \mathbb{R}$. Therefore

$$\|u'_{\varepsilon_j}(s)\|^2 \geq \frac{2E\|q_b - q_a\|^2}{2(E - V(u_{\varepsilon_j}(s)/\varepsilon_j))} \geq \|q_b - q_a\|^2. \tag{3.122}$$

It follows that

$$\|t'_{\varepsilon_j}(s)\| \geq \frac{\|q_b - q_a\|}{\sqrt{2E}}$$

and hence by (3.119)

$$t_{\varepsilon_j}(\tau_2) - t_{\varepsilon_j}(\tau_1) \geq \frac{(\tau_2 - \tau_1)\|q_b - q_a\|}{\sqrt{2E}}. \tag{3.123}$$

Passing to the convergence subsequence corresponding to t_0, and taking the limit as

$j \to \infty$ gives that

$$t_0(\tau_2) - t_0(\tau_1) \geq \frac{(\tau_2 - \tau_1)\|q_b - q_a\|}{\sqrt{2E}} > 0. \tag{3.124}$$

Hence t_0 is strictly monotonic. □

It is possible that $\cap_{j=1}^{\infty} t_{\varepsilon_j}((0,1)) = \emptyset$, in this case there exists a further subsequence of $\{\varepsilon_j\}_{j=1}^{\infty}$ such that $T_{\varepsilon_{j_k}} \to 0$ as $k \to \infty$. The following lemma shows that this is in fact impossible, and therefore as $\varepsilon \to 0$ the particle cannot arrive 'instantly' at it's destination.

Lemma 3.69. *It holds that* $\inf_j T_{\varepsilon_j} \geq C > 0$.

Proof. First note that since $\max V = 0$ it holds that

$$\frac{1}{\sqrt{2(E - V(u_{\varepsilon_j}(s)/\varepsilon))}} \geq \frac{1}{\sqrt{2E}}. \tag{3.125}$$

Furthermore, proceeding as in the proof of lemma 3.64,

$$2E\|q_b - q_a\|^2 \leq G_\varepsilon(u_{\varepsilon_j}) \leq 2\left(E + \max_{[0,1]^d}\{-V(x)\}\right)\|q_b - q_a\|^2. \tag{3.126}$$

As u_ε is a sequence of geodesics parameterised by arc length it follows that $G_{\varepsilon_j}(u_{\varepsilon_j}) = 2(E - V(u_{\varepsilon_j}(s)/\varepsilon_j))\|u'_{\varepsilon_j}(s)\|^2 \in \mathbb{R}$. Therefore

$$\|u'_{\varepsilon_j}(s)\|^2 \geq \frac{2E\|q_b - q_a\|^2}{2(E - V(u_{\varepsilon_j}(s)/\varepsilon_j))} \geq \|q_b - q_a\|^2. \tag{3.127}$$

Combining equations (3.125) and (3.127) gives that

$$T_{\varepsilon_j} = \int_0^1 \frac{\|u'_{\varepsilon_j}(\tau)\|}{\sqrt{2(E - V(u_{\varepsilon_j}(\tau)/\varepsilon))}} \mathrm{d}\tau \geq \frac{\|q_b - q_a\|}{\sqrt{2E}}.$$

Taking the infimum over j proves the result. □

It is interesting to observe the last inequality in the proof of 3.69, which provides an estimate on the arrival time in terms of both the energy imposed and the distance between those states.

With lemma 3.69 it is possible to define $\bar{T} := \inf_j T_{\varepsilon_j} > 0$ and hence $\cap_{j=1}^{\infty} \mathrm{dom}(t_{\varepsilon_j}) = [0, \bar{T}] \neq \emptyset$. Applying theorem 3.67 in combination with lemmas 3.68 and 3.69 we conclude the following result.

Theorem 3.70. *The sequence of inverses* $\{t_{\varepsilon_j}^{-1}\}_{j=1}^{\infty}$ *is precompact on* $[0, \bar{T})$.

Proof. Applying theorem 3.67 in combination with lemmas 3.68 and 3.69 gives that $\{t_{\varepsilon_j}^{-1}\}_{j=1}^{\infty}$ converge uniformly on $(0, \bar{T})$. Since $t_{\varepsilon_j}^{-1}(0) = 0$ for all j, the result follows □

The following corollary is immediate.

Corollary 3.71. *The sequence of trajectories converge uniformly on* $(0, \bar{T})$.

As the arrival time at q_b varies with ε the information in corollary is the best attainable. In the literature, such as in [LPV88, BD98], where the authors consider the homogenisation of the action functional, the authors prescribe the arrival time along with the start and end points. The consequence of this is that the total energy E of the solution is adjusted with respect to ε, and, as the action is not explicitly dependant on E, classical homogenisation results for integral functionals apply. In the case of the Jacobi metric, where the functional explicitly depends on E, we have two options. The first option is to fix the total energy independent of ε, as we have done here, then as we have seen issues arise with the convergence of trajectories on different domains, but partial information about the effective behaviour can be recovered. The second option is to fix the arrival time. The problem with this is that the energy E would have to change with ε. Provided we could establish enough information about how E depends on ε, we would then have to try to homogenise a sequence of functionals whose integrand depends on ε. This proves to be particularly difficult as we have already seen in section 3.3, and there is no general theory that currently applies to these situations. It would prove to be an interesting problem to see how to determine the Γ-limit of such functionals and we reserve this for future investigation.

Let us discuss the motion of the particle when $0 \geq E \geq \min V$. In these cases the Jacobi metric corresponds to a degenerate Riemannian manifold. Homogenisation of such length functionals is again open for future investigation. We expect that, as $\varepsilon \to 0$, that the geodesic eventually touches the region where the metric degenerates, which corresponds the the particle taking an infinite amount of time to reach a state different from the end point. This would indicate that for the corresponding equation for Newton's second law has no solution to the boundary value problem. Studying the homogenisation of degenerate Riemannian length functionals may still be of interest for other applications.

3.5.3 A Model of Hamiltonian Dynamics in Discontinuous Potentials

When interpreting the results of this chapter in the context of Hamiltonian dynamics, one has to exercise caution. Here we have assumed that the motion of a Hamiltonian potential is directly modelled using the Maupertuis principle. Whereas the typical route

is to study the motion as the minimiser of the action functional. In the case that $V \in L^\infty(\mathbb{R}^d)$, the results of section 3.3 could still be applied to describing effective dynamics under V. Such potentials can still be considered as physically reasonable, as such discontinuities could represent physical constraints on where the particle may go. As V is sufficiently irregular, equation (3.102) is not well defined. In addition, as the Maupertuis principle is based on the equivalence of Euler-Lagrange equations, which in turn require that V is differentiable almost everywhere. Therefore we are in the following position, we have two principles that describe the motion of a Hamiltonian particle, that are equivalent when V is differentiable. When $V \in L^\infty(\mathbb{R}^d)$, it is no longer possible to show that these formulations are equivalent. However, there are no immediate physical reasons why one principle should be favoured over another in this context. We therefore make the assumption, that the Maupertuis principle is the correct functional to describe the motion of a Hamiltonian particle in a $L^\infty(\mathbb{R}^d)$ potential. It would prove to be very interesting to study the Γ-convergence of the corresponding action functionals and to see if similar results to those in section 3.3 hold. This is left open for further investigation.

Chapter 4

Molecular Simulations via the Maupertuis Principle

4.1 Introduction

The purpose of this chapter is to study consistent numerical methods to find the minimum of the functional

$$L[u] = \int_0^1 a(u(\tau))\|u'(\tau)\|d\tau$$

over the class of Lipschitz curves joining q_a to q_b in $\mathbb{R}^d$. Here we will assume that $0 < a \leq \beta$, for some $\beta \in \mathbb{R}$, we also assume that $a \in C^2(\mathbb{R}^d)$. The main application of this we have in mind, as previously discussed in subsection 3.1.2, is to compute solutions of

$$q''(t) = -\nabla V(q(t)), \tag{4.1}$$

subject to $q(0) = \xi_1$ and $q(T) = \xi_2$ for some T to be determined, using the Maupertuis principle. As discussed in subsection 3.1.2, the solutions of (4.1) correspond to critical points of the functional

$$\int_0^1 \sqrt{2(E - V(u(\tau)))}\|u'(\tau)\|\mathrm{d}\tau$$

when time is reparameterised by

$$t(s) = \int_0^s \frac{\|u'(\tau)\|}{\sqrt{2(E - V(u(\tau)))}}\mathrm{d}\tau. \tag{4.2}$$

It's clear that in order to find a solution of (4.1) we need to prescribe the energy along the trajectory in advance. The value for T is connected to the value for E we choose as can be seen in equation (4.2), noting that $T = t(1)$. The aim of this chapter is

to build on the existing results in [SZ09a, SZb, SZa]. In [SZ09a, SZb, SZa] the authors develop a curve shortening procedure to numerically find geodesics based on the Birkhoff method [Bir66], which we outline below. The Birkhoff method provides a technique to compute extended geodesics using only short geodesic segments. In this context we say that a geodesic segment is short if it has end points which are sufficiently close together. Specifically, in our setting we say that the end points, q_a and q_b, are sufficiently close together if

$$\|\xi_1 - \xi_2\| \leq \frac{1}{6(4 + 5\|\nabla a\|_\infty)\sqrt{d-1}}.$$

This definition of points being sufficiently close follows from [SZb, Definition 4.2]. In [SZb] it is demonstrated, under suitable regularity assumptions and boundedness conditions on the metric, that when the end points are sufficiently close then there exists a length minimising geodesic joining them. Throughout this chapter we will refer to geodesics with sufficiently close endpoints as local geodesics. Otherwise they will be referred to as global geodesics. The terms global geodesic and extended geodesic are synonymous.

We will now describe Birkhoff's method to compute an extended geodesic. The initial step is to join q_a to q_b with a curve, which serves as an initial guess. Should such a curve fail to exist, then it's clear that the minimum problem doesn't have a solution. The initial curve is then divided into sufficiently small segments. The endpoints are then enumerated in sequence. To reduce length we move the nodes in turn, first moving the odd numbered nodes, and then the even numbered nodes, ignoring the endpoints. While we specify an order in which to move the nodes, other orders may be used. When node i is selected, compute the geodesic joining the nodes $i-1$ and $i+1$ and determine the midpoint along this geodesic. We use the midpoint along the geodesic as the new position for node i. This process is repeated ad infinitum and is shown to converge to a geodesic joining the endpoints q_a, q_b. We will further illustrate this process with an example. Consider the case of a Euclidean metric, where we know the geodesics between any two points to be the straight line segment connecting them.

The first step of the Birkhoff method applied to the Euclidean metric is illustrated in figure 4-1. We know that the geodesic joining q_a to q_b is the straight line segment. The initial guess is the dashed line, we then select the grey nodes and compute the geodesics between them, which are straight lines. We then update the positions of the black nodes to the midpoints along these lines, this produces the solid curve. In the next turn we then select the black nodes and compute the geodesics joining these points, updating the positions of the grey nodes as the midpoints along these geodesics, producing the dotted line. This procedure continues and the subsequent approximations converge to

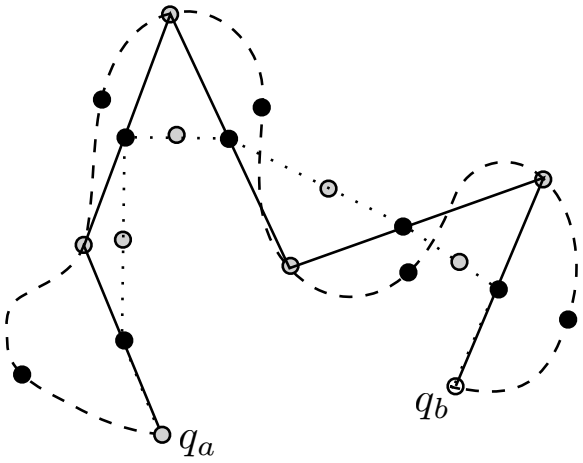

Figure 4-1: *Birkhoff Method Applied to the Euclidean Metric. Reproduced with permission from [SZb].*

the geodesic joining q_a to q_b.

The Birkhoff method is the foundation on which the numerical methods in [SZ09a, SZb, SZa] are based. Any numerical implementation of the Birkhoff procedure that will compute extended geodesics needs two key features; a global procedure that handles the movement of the node positions and a local procedure that computes the geodesics between the nodes and returns the midpoint. There are benefits to this separation between local and global calculations; we can make additional a priori assumptions on the local procedure to ensure convergence and efficiency. A constructive method, to compute local geodesics, was devised and analysed in [SZb]. A full description of the procedure developed in [SZb] is given in subsection 4.3.2. We also study a simplification of the method which we describe in subsection 4.3.1. An alternative approach in [SZ09a], which may be applied to calculating local geodesics, based on a gradient descent method, is described in subsection 4.3.3 from.

To the best of the authors knowledge, the first time curve shortening techniques are applied to Riemannian length functionals is in [SZ09a, SZb]. Curve shortening methods have been used in molecular dynamics simulations before, however they are typically applied to discretisations of the action functional [FS02, GW92]. We mention that molecular simulation via the Maupertuis principle has been studied before in [BA90]. The approach in [BA90] performs a global minimisation of the length functional over a linear combination of basis functions; in contrast to the approach taken here, and in [SZ09a, SZb], where the global minimisation problem is replaced with multiple local minimisation problems.

We will develop an approach that combines the features of [SZ09a, SZb] approaches in an attempt to speed up the calculations without losing the quality of the answers. Furthermore we will also investigate the effects of parallelising this procedure. The nature of the Birkhoff procedure makes it easy implement in parallel with little modification. In particular we perform a detailed numerical investigation into its efficiency, and apply the method to the solution of a molecular dynamics problem. Specifically, we perform a simulation of a single butane molecule as it changes its conformation using the model in [Fis97].

For each procedure, we first analyse it mathematically, followed by a detailed description of the implementation of the procedure along with its pseudocode. With each procedure, where appropriate, we will also compare the respective performance on an example problem for which an exact solution is known.

In the pseudocode we make the following definitions. The function `Norm` takes a single double precision array of length d, representing a vector in $\mathbb{R}^d$, and returns the Euclidean norm of the vector. The functions `Abs` and `Sign` take a double precision number and return the absolute value and sign of the number respectively.

4.1.1 Results of this Chapter

The key contributions of this chapter towards curve shortening methods for molecular simulations are as follows.

In section 4.4 we develop a 'predictor-corrector' technique for calculating local geodesics. The aim was to take the method from [SZb] and include measures that would typically improve the time it took to complete a calculation. The original procedure from [SZb], proven to be numerically consistent, is discussed in subsection 4.3.2, with our modifications detailed in section 4.4. The success of these modifications, tested on a particular example, is analysed in section 4.5. We also determine the performance of the procedure as d increases; specifically we find that as the amblent d increases, the new local procedure is faster in relation to the method in [SZb].

In section 4.6 we discuss the parallelisation of the global procedure to compute extended geodesics. This extends the work of [SZa] by combining the faster local procedure in section 4.4 with parallel computing techniques. The performance of the implementation is again measured on an example problem. In particular, we study the effect that changing the number of cores has on the performance of the procedure. We also study the performance of the procedure for higher dimensional problems.

In section 4.7 we apply the code to the simulation of the change in conformation of butane. This problem was considered in [Fis97]. In [Fis97] the statistics for the change in conformation were studied and simulated. In section 4.7 we compute single

low energy trajectories for the transition and study the motion along the resulting curve. In particular we are able to quantify at which energies the motion of the butane molecule is nontrivial. The notion of nontriviality being that the resulting solution fails to be close to the straight line joining the initial and final configurations.

4.2 The Global Procedure

The global procedure handles the motion of the endpoints of the geodesic segments and passes the computation of local geodesics between nodes to the procedures that we will develop in 4.3. The procedure mirrors the Birkhoff method, with the computation of the local geodesics being performed by `LocalProcedure`. The mathematical description of this process was discussed in the introduction of this chapter. Additional measures are included to ensure that the procedure terminates at a reasonable point and are outlined below.

Discussion of Pseudocode for the Global Procedure

The pseudocode for the process is described in procedure 4.1. The process first initialises necessary variables and assigns the order for the nodes to be moved in $NNumber$. The variable Tol measures the average distance travelled by each node and is used to determine when the procedure should stop. The while statement in line 3 measures the movement of the nodes against a specified tolerance. As the while statement executes nodes are selected according to the order described in $NNumber$. The node is moved to the midpoint of the geodesic joining the neighbouring nodes using `LocalProcedure`, which serves as a placeholder for any of the procedures discussed in 4.3.

Before studying the implementation of the procedure, we need to determine how to numerically compute the local geodesics. This is performed in the next section. We will also modify the global procedure to run in a parallel architecture, this is detailed in section 4.6.

4.3 Local Procedure

In this section we describe a new procedure for determining the shortest distance between two points, assuming that the end points are sufficiently close together. Our approach is a mixture of two approaches that have been previously developed by [SZ09a, SZb]. We describe the approaches of [SZ09a, SZb] in detail before discussing our predictor-corrector approach.

Procedure 4.1: Global Curve Shortening Procedure

Inputs : CurveData $\in \mathcal{M}^{(2N+1)\times d}(\mathbb{R})$: Each row representing a d-dimensional vector corresponding to a position along the initial curve.
TolMin $\in (0,1)$: Sets stopping condition for the procedure.

Output: CurveData $\in \mathcal{M}^{(2N+1)\times d}(\mathbb{R})$: Each row representing a d-dimensional vector corresponding to a position along the calculated curve.

```
Tol ← 1;
NNumber ← {3, 5, ..., 2N − 1} ∪ {2, 4, ..., 2N};
while Tol > TolMin do
    for i ← 1 to N − 2 do
        I = NNumber(i);
        NPosition ← LocalProcedure (CurveData(I-1),
        CurveData(I+1),Params) ;
        Tol = Tol + Norm (NPosition − CurveData(I));
        CurveData(I) = NPosition;
    end
    Tol = Tol ·N;
end
```

4.3.1 A Step Based Approach

A Single Step Procedure

The approach taken is similar to the Birkhoff method. However, the nodes are moved in the space orthogonal to the tangent of the initial curve to reduce length. A core assumption is that the endpoints, denoted as q_a, q_b, are sufficiently close together. This ensures that the geodesic can be expressed as a graph over the straight line segment joining q_a to q_b [SZb]. We then partition the straight line joining q_a to q_b into $2N+1$ pieces of equal length and enumerate these points as $\{q_1 = q_a, q_2, ..., q_N = q_b\}$. Also, suppose that N is selected large enough to ensure that $\varepsilon := \|q_2 - q_1\| < 1$. Define the linear space $\mathcal{N} := \{x \in \mathbb{R}^d : x \cdot (q_b - q_a) = 0\}$ and let $\mathcal{B} := \{e_1, ...e_{d-1}\}$ be an orthonormal basis of $\mathcal{N}$, this describes a hyperplane that is orthogonal to the line joining q_a to q_b. Then, for each $i \in \{2, ..., N-1\}$ and $j \in \{1, ..., d-1\}$ define $q_{i,j}(\delta) = q_i + \delta e_j$, with $\delta < \varepsilon^{2+\alpha}$ for some $\alpha \in (\frac{1}{2}, 1)$. Furthermore, let $q_{i,-j}(\delta) = q_i - \delta e_j$ and $q_{i,0}(\delta) = q_i$. We define the approximate Riemannian distance between two points q_a and q_b by

$$L(q_a, q_b) := \left(\frac{a(q_a) + a(q_b)}{2}\right) \|q_b - q_a\|. \tag{4.3}$$

Then the procedure is as follows, firstly for each $i \in \{3, 5, \dots, 2N-1\}$ determine j such that

$$\min_{\pm j \in \{0,1,\dots,d-1\}} L(q_{i-1}, q_{i,j}(\delta)) + L(q_{i,j}(\delta), q_{i+1}),$$

is attained. When j is found, update $q_{i,j} \to q_i$. Then proceed to the next odd indexed node, finally repeat the process for $i \in \{2, 4, \dots, 2N\}$. Should $j = 0$ for all interior $i \in \{2, \dots, 2N\}$ the procedure terminates and the final node positions returned, otherwise return to the minimisation problem for $i \in \{3, 5, \dots, 2N-1\}$. The reason for choosing δ and α as outlined above is to ensure the convergence of the procedure, guaranteed for sufficiently regular a and close end points by the theory in [SZb].

Discussion of the Pseudocode for the Single Step Procedure

To describe the procedure in detail we provide a pseudocode. Within the pseudocode we use particular functions whose behaviour we will describe now. The function `Len` takes three double precision arrays of length d representing vectors, which we denote mathematically as q_a, q_b and q_c, and returns the value

$$L(q_a, q_b) + L(q_b, q_a),$$

where L is defined in (4.6). The function `FindNormalBasis` takes a single double precision array and returns a $d \times d$ array. The columns of this array represent an orthonormal basis of $\mathbb{R}^d$ which includes the normalised input vector. We do not discuss this process in detail here but include our implemented versions in appendices A and B.

The single step procedure starts with procedure 4.2. The purpose of this particular procedure is to provide a mechanism for selecting nodes in a specified order; that is, odd numbered nodes and then even numbered nodes. Other node orders may be used without affecting the consistency of the method, the effect on performance for doing this is not studied here. It also contains the trigger to halt the procedure. The procedure is stopped when no nodes can be moved to further reduce length; in this case the flag $NewMin$ remains zero at the end of the iteration and the procedure stops. By returning to procedure 4.2 we would either select the next node or terminate the procedure; this is determined by the state of the flag $NewMin$ and the value of i as described above. The performance of this procedure is measured against the other local methods in section 4.5.

Once line 2 of procedure 4.2 is reached procedure 4.3 is called, this procedure simply uses the function `FindNormalBasis` to determine a basis for $\mathcal{N}$ and stores it ready for the node movement procedure. The process then resumes 4.2 to select the nodes for

Procedure 4.2: Local Node Handling for Curve Shortening Procedures

Inputs : $\mathsf{Node} \in \mathcal{M}^{N\times d}(\mathbb{R})$: Each column representing a d-dimensional vector corresponding to a position along the initial curve. It will always hold that $\mathsf{Node}(1) = q_a$ and $\mathsf{Node}(N) = q_b$.
$\delta \in (0, 1)$: Selected as described in the introductory text (step method only).

Output: $\mathsf{Node} \in \mathcal{M}^{N\times d}(\mathbb{R})$: Each column representing a d-dimensional vector corresponding to a position along the calculated curve. It will always hold that $\mathsf{Node}(1) = q_a$ and $\mathsf{Node}(N) = q_b$.

```
1  NodeOrder ← {3, 5, ..., 2N − 1} ∪ {2, 4, ..., 2N};
2  ⟨ Computation of 𝒩 for Single/Multi Step Code Only — Procedure: 4.3 ⟩;
3  DirectionOrder ← {−(d − 1), ..., −1} ∪ {1, ..., (d − 1)};
4  NewMin ← 1;
5  while NewMin ≠ 0 do
6  |  NewMin ← 0;
7  |  for i ← 1 to N − 2 do
8  |  |  ⟨ Node Movement — Procedure: Either 4.4 (Steps) or 4.7 (Gradient) ⟩;
9  |  end
10 end
11 ReturnMidpoint(Node);
```

movement.

Procedure 4.3: Procedure to Determine the Basis of Normal Directions to the Initial Line in the Step Procedures

```
1  ⟨ Called in Line 2, Procedure: 4.2 ⟩;
2  Tangent ← q_b − q_a;
3  {e_1, ..., e_{d−1}} ← FindNormalBasis(Tangent) ;
4  ⟨ Return to Line 2, Procedure: 4.2 ⟩;
```

At line 8 of procedure 4.2, we have selected node I and are ready to move it in the directions of $\{e_1, \ldots, e_{d-1}\}$ to attempt to reduce length. For the single step procedure this involves moving the node a 'single step' in each basis direction of $\mathcal{N}$ to test for length reduction; the size of the step is determined by δ. The movement of the nodes is handled by procedure 4.4. A significant portion of the procedure is dedicated to calculating the direction that produces the greatest length reduction. In line 7 the node is moved a δ-step in the selected basis direction; the selection of the current basis direction is handled by the variable j. Should there be a direction where the length is reduced then the $NewMin$ flag is activated to indicate that neighbouring nodes need to be checked for length reduction; this flag is set in line 9. Before returning to procedure 4.2, and subsequently moving to the next node, it is essential we record this new node position

which is performed by 4.5.

Procedure 4.4: Procedure to Determine the Direction that Minimises Length for the Step Procedures

```
1  ⟨ Called in Line 8, Procedure: 4.2 ⟩;
2  ⟨ Input: NodeOrder, DirectionOrder, NewMin, Node ⟩;
3  MinJ ← 0, I ← NodeOrder(i);
4  BLen ← Len(Node(I-1),Node(I),Node(I+1));
5  for j ← 1 to 2(d − 1) do
6  |  J ← Abs(DirectionOrder(j)), sJ ← Sign(DirectionOrder(j));
7  |  TLen ← Len(Node(I-1),Node(I)+sMinJ · δe_J,Node(I+1));
8  |  if TLen < BLen then
9  |  |  NewMin ← 1, BLen ← TLen, MinJ ← J, sMinJ ← sJ;
10 |  end
11 end
12 ⟨ Return to Line 1, Procedure: 4.5 (Single Step) or 4.6 (Multiple Step) ⟩;
```

Once the optimal node position is determined, it is recorded by 4.5. For the single step procedure this is performed by updating the position in the *Nodes* array and then returning to line 8 of procedure 4.2.

4.3.2 A Multiple Step Procedure

We now discuss an extension of the single step procedure with a view to improving performance. This extension results in the procedure described in [SZb]. The procedure is structurally similar to the single step method. The difference is that before updating the node position, we check whether moving additional δ-steps in the optimal length reducing direction would further decrease length. By performing this check we can reduce the number of times that we search for the optimal direction, which could improve performance for problems in high dimensions.

Mathematically, we begin by repeating the single step procedure. The difference occurs when the optimal direction $j \in \{-(d-1), \ldots, d-1\}$ for node $i \in \{1, \ldots, N\}$ has

Procedure 4.5: Node Movement Procedure for Single Step Procedure

```
1  ⟨ Called in Line 12, Procedure: 4.4 ⟩;
2  ⟨ Input: I, Node, MinJ, sMinJ ⟩;
3  if MinJ ≠ 0 then
4  |  Node(I) ← Node(I)+sMinJ · δe_MinJ;
5  end
6  ⟨ Return to Line 8, Procedure: 4.2 ⟩;
```

been determined. For the single step method we would have previously set $q_{i,j}(\delta) \to q_i$. Instead of this, we test whether $q_{i,j}(2\delta)$ reduces the length further. If it reduces the length further then we proceed to test $q_{i,j}(3\delta)$ and continue increasing the number of δ-steps in this way until the length is no longer reduced. Finally we update the node position $q_{i,j}(N\delta) \to q_i$, where N is the last length reducing multiple of δ and continue on to moving the next node.

Discussion of the Pseudocode for the Multiple Step Procedure

As before we discuss the pseudocode for the multiple step curve shortening procedure. The procedure begins in an identical fashion to the single step code, however instead of using procedure 4.5 to update the node position, we use procedure 4.6. The replacement procedure provides the mechanism that will continue to test for length minimisation in the previously determined optimal direction. It is essentially a linear search in N, starting with $N = 2$, for which $q_{i,j}(N\delta)$ minimises the length the most. The process is guaranteed to terminate in a finite number of steps by [SZb, Proposition 3.5].

Procedure 4.6: Node Movement Procedure for Multiple Step Procedure

```
1  ⟨ Called in Line 12, Procedure: 4.4 ⟩;
2  ⟨ Input: I,BLen, Node, sMinJ, MinJ ⟩;
3  if MinJ ≠ 0 then
4  |  TLen ← 0, Steps ← 1;
5  |  while TLen < BLen do
6  |  |  BLen ← TLen;
7  |  |  Steps = Steps + 1;
8  |  |  TLen ← Len(Node(I-1),Node(I) + sMinJ · Steps · e_MinJ,Node(I+1));
9  |  end
10 |  Node(I) ← Node(I) + sMinJ · (Steps − 1) · e_MinJ;
11 end
12 ⟨ Return to Line 8, Procedure: 4.2 ⟩;
```

The performance of this procedure is measured against the other local methods in section 4.5.

4.3.3 A Gradient Descent Approach

This next section is dedicated to the discussion of a gradient descent approach in [SZ09a]. For this procedure we do not require the assumption that the geodesic is a graph over the initial curve. As before, we partition the straight line joining q_a to q_b into $2N + 1$ pieces of equal length and retain the same notation as for the previous local procedures. We take N large enough to ensure that $\|q_2 - q_1\| < 1$.

The gradient descent curve shortening procedure is as follows. Select a node with index $i \in \{3, 5, \dots, 2N-1\}$, for this node compute the discrete tangent $\tau_i \in \mathbb{R}^d$, for the curve, as

$$\tau_i = \frac{1}{2}\left((q_{i+1} - q_i) - (q_i - q_{i-1})\right).$$

With the vector τ_i define the space $\mathcal{N}(\tau_i) := \{x \in \mathbb{R}^d \colon x \cdot \tau_i = 0\}$. Let $\mathcal{B}(\tau_i) := \{e_1, \dots e_{d-1}\}$ be a basis for $\mathcal{N}(\tau_i)$. Note that we drop the dependance of e_j with τ_i for notational simplicity, and set $e_{-j} = -e_j$. Set $\varepsilon = \|\tau_i\|/4$ and for a small fixed number Δs set $\Delta\xi = 2\Delta s/\varepsilon^2$. The purpose of the parameter Δs will be described below. We retain the definition $q_{i,j}(\delta)$ as in the previous local procedures. Fix $j \in \{-(d-1), \dots, -1, 1, \dots, d-1\}$ and compute the following quantities

$$\begin{aligned} L^+ &= L(q_{i-1}, q_{i,j}(\varepsilon)) + L(q_{i,j}(\varepsilon), q_{i+1}) \\ L^- &= L(q_{i-1}, q_{i,-j}(\varepsilon)) + L(q_{i,-j}(\varepsilon), q_{i+1}). \end{aligned}$$

With L^+ and L^- we can estimate a new position for q_i to test for length reduction. The candidate for a new node position is determined by

$$q_{i,j}\left(-\Delta\xi(L^+ - L^-)\right) \equiv q_i - \Delta\xi(L^+ - L^-)e_j.$$

With this information it is then possible to compute

$$L(q_{i-1}, q_{i,j}(-\Delta\xi(L^+ - L^-))) + L(q_{i,j}(-\Delta\xi(L^+ - L^-)), q_{i+1})$$

and therefore determine

$$\min_{\pm j \in \{0,1,\dots,d-1\}} L(q_{i-1}, q_{i,j}(\delta)) + L(q_{i,j}(\delta), q_{i+1})$$

and its minimum at j. Then update $q_{i,j} \to q_i$. Then the process continues to the next odd indexed node and then for $i \in \{2, 4, \dots, 2N\}$. Should $j = 0$ for all $i \in \{2, \dots, 2N\}$ the procedure terminates and the final node positions returned; otherwise we move to the first odd numbered node and repeat the above.

Comparison with Step Based Methods

Let us discuss how this approach compares with the step based local procedures. The first observation is that with this procedure the curve is no longer confined to a grid. This is owing to the fact that the basis of normal directions to the curve is determined at each node. With the step procedures it is necessary to either make a single small step, or by

a process of elimination, determine how far one should shift the node to determine if the length is reduced. In the gradient approach the distance to travel in a normal direction is given explicitly. Note that the proposed distance is to test for length reduction. The consequence of this is that while the procedure has similar mechanics to the single step procedure, the possibility of making larger steps in the optimal direction is presented, a clear computational advantage. The limitations of this method as opposed to the step methods are the absence of rigourous consistency results. The method is based upon the discretisation of a parabolic equation, it is essential that artificial time parameter Δs is chosen sufficiently small, otherwise the process may diverge or oscillate. The quantity $\Delta\xi$ in this context may be interpreted as the Courant-Friedrichs-Lewy number for this discretisation. The step methods present the advantage that minimising curves on a prescribed grid to in fact converge, and this is demonstrated rigourously [SZb]. For these reasons we do not include a comparison with the other local methods, but rather include these details to explore the literature.

Discussion of the Pseudocode for the Gradient Descent Procedure

The procedure begins as the step methods with procedure 4.2. The first difference, in comparison with the step methods, is that as we no longer use a universal basis of normal directions and hence we omit line 2 from this process. Once the procedure is in a position to move a node, we use a different process to compute the new node position, this is described in procedure 4.7.

In procedure 4.7 the first calculations are to determine the discrete tangent and a basis of normal directions, this is achieved in lines 5 and 6. Once the basis of normal directions is established it remains to compute the new node position for each direction and test whether this reduces length. The test for optimal length reduction and recording the optimal directions is exactly as in the step methods and performed in lines 13 and 14. Additional information needs to be recorded if optimal direction is found, namely the corresponding values for L^+ and L^-, in order to reconstruct the optimal position for the node update in procedure 4.8. We note also that procedure 4.7 also controls the flag $NewMin$ to determine if the process should continue.

As in the step methods, if there is no better position for the node, then the process moves to the next node or terminates as appropriate. If there is a length minimising direction then the required step distance is reconstructed from the stored information and the node position updated.

Procedure 4.7: Procedure to Determine the Direction that Minimises Length for the Gradient Descent Procedure.

```
1  ⟨ Called in line 8, Procedure: 4.2 ⟩;
2  ⟨ Input: NodeOrder, DirectionOrder, NewMin, Node, Δs ⟩;
3  MinJ ← 0, I ← NodeOrder(i);
4  BLen ← Len(Node(I-1),Node(I),Node(I+1));
5  Tangent ← ½(Node(I+1) + Node(I-1));
6  {e_1, ..., e_{d-1}} ← FindNormalBasis(Tangent) ;
7  ε ← Norm(Tangent)/4, StepSize ← 2Δs/ε²;
8  for j ← 1 to 2(d − 1) do
9  |  J ← Abs(DirectionOrder(j)), sJ ← Sign(DirectionOrder(j));
10 |  FLen ← Len(Node(I-1),Node(I)+εe_j,Node(I+1));
11 |  RLen ← Len(Node(I-1),Node(I)−εe_j,Node(I+1));
12 |  StepDistance ← −sJ · StepSize · (FLen − RLen);
13 |  TLen ← Len(Node(I-1),Node(I)+StepDistance · e_j,Node(I+1));
14 |  if TLen < BLen then
15 |  |  MinJ ← j, BLen ← TLen, BFLen ← FLen, BRLen ← RLen;
16 |  |  sMinJ ← sJ, NewMin ← 1;
17 |  end
18 end
19 ⟨ Return to line 1, Procedure: 4.8 ⟩;
```

Procedure 4.8: Node Movement Procedure for Gradient Flow Procedure.

```
1  ⟨ Called in line 8, Procedure: 4.2 ⟩;
2  ⟨ Input: MinJ, sMinJ, StepSize, BFLen, BRLen, Node, I ⟩;
3  if MinJ ≠ 0 then
4  |  StepDistance ← −sMinJ · StepSize · (BFLen − BRLen);
5  |  Node(I) ← Node(I)+StepDistance · e_MinJ;
6  end
7  ⟨ Return to line 8, Procedure: 4.2 ⟩;
```

4.4 A Predictor-Corrector Local Method

In this section we propose a new local curve shortening procedure. The new approach is an attempt to unite the consistency and convergence of the step procedures with the efficiency of having a formulaic description for the new node positions. We attempt to predict the new node position by calculating the gradient of the length functional with respect to displacements of the selected node. The prediction is refined by testing further displacements in a manner similar to a binary search. All of the predictions and corrections are performed in relation to the grid from the step methods, to ensure consistency and length reduction.

In the literature there exist other techniques to solve nonlinear optimisation problems, such as quasi-Newton methods [BGLS06, Section 4.4] and the BFGS method [BGLS06, Section 4.7]. These approaches could also be used to determine the optimal position of a given node. The issue here is that these techniques rely on the approximation of the Hessian matrix of the length functional which for high dimensional problems becomes expensive. The Birkhoff method discards most of the previous optimisation steps in its iterations, consequently we would require making several of these expensive calculations at every step. The predictor-corrector method we develop here aims to form a compromise between accuracy and efficiency.

The idea of this new procedure is to follow the structure of the single step procedure; the size of the step determined by an approximation of the gradient of the length functional. We select the nodes in turn and will proceed to move them in a universal basis of normal directions to the initial curve. Retaining the same notation as for the step methods, choose $\alpha \in (\frac{1}{2}, 1)$ and $\delta < \varepsilon^{2+\alpha}$. As with the other step methods, the nodes are moved by vectors in

$$\mathcal{G} := \{m\delta e_j \colon m \in \mathbb{N}, j \in \{-(d-1), \dots, d-1\},$$

the set $\mathcal{G}$ indirectly defines vertices of the grid where we may position our nodes. Rather than making a single step ($m = 1$), or performing a linear search in m, we predict an optimal m and attempt to perform a more efficient search. The prediction will be based on calculating the gradient of the length functional with respect to displacements by $\mathcal{G}$. Consequently, we combine the consistency of the step method, with the potential of speeding up the process with a formulaic approach.

Let us describe the procedure mathematically before discussing the pseudocode. Given the end points of the curve q_a and q_b we calculate the tangent to the initial curve, which is $q_b - q_a$. As with the step methods, we compute the space of normal directions, that is $\mathcal{N} := \{x \in \mathbb{R}^d \colon x \cdot (q_b - q_a) = 0\}$ and calculate a basis for $\mathcal{N}$. Denote

the basis for $\mathcal{N}$ by $\{e_1, \dots, e_{d-1}\}$. Furthermore let $A(\mathcal{N})$ be the change of basis mapping between $\{(q_b - q_a)/\|q_b - q_a\|, \dots, e_{d-1}\}$ and the standard basis of $\mathbb{R}^d$. We will proceed by considering each node in turn, first by those with odd indices, followed by those with even indices. Suppose that we have $i \in \{3, 5, \dots, 2N-1\}$ fixed and are in a position to move node q_i in order to reduce length. Then define the function

$$\mathcal{L}(\eta) := L(q_{i-1}, q_i + \eta) + L(q_i + \eta, q_{i-1})$$

and consider its gradient $\nabla\mathcal{L}$. Clearly, the best direction to test for length reduction is $-\nabla\mathcal{L}(0)$ as this represents the direction of maximal decrease. Define $\varepsilon = \|q_2 - q_1\|$ and fix $\alpha \in (\frac{1}{2}, 1)$, then choose $\delta < \varepsilon^{2+\alpha}$ to ensure that any variations of the curve in δ-steps are convergent, as with the step method. We discretise the computation of this gradient, for the jth, with $j \in \{1, \dots, d-1\}$, component we set

$$D_{j+1}\mathcal{L} := \frac{1}{2\delta\varepsilon}\left(L(q_{i-1}, q_i + \delta e_j) + L(q_i + \delta e_j, q_{i-1}) - L(q_{i-1}, q_i - \delta e_j) - L(q_i - \delta e_j, q_{i-1})\right).$$

We set $D_1\mathcal{L} = 0$ to ensure that the nodes move only in normal directions to the initial curve. The quantity D_j converges by [SZb, Lemma 3.4]. During the calculation of $D_j\mathcal{L}$ there is sufficient information to determine whether there exists a single δ-step movement that will minimise length. Should a single δ-step not exist, then the process proceeds to the next node.

In general we expect the direction suggested by $\mathcal{D}\mathcal{L} := \{D_j\mathcal{L}\}_j$ not to lie in a set of the form $\{\sum_{j=1}^{d-1} m_j e_j : m_j \in \mathbb{Z}, j \in \{1, \dots, d-1\}\}$, and hence does not automatically lie on the grid of the step methods. In order to have the consistency of the step methods we require the shifted nodes to remain on the grid. Let us outline how we achieve this.

In order to ensure that shifting the node in the gradient direction reduces length we rescale $\mathcal{D}\mathcal{L}$ by $C\varepsilon^2$, where C is a problem dependent constant. The next step of the procedure is to now shift the currently selected node in multiples of this gradient distance. To do this, rather than move in multiples of $C\varepsilon^2\mathcal{D}\mathcal{L}$ we consider the following mapping

$$S(\tau) = \left(\frac{\mathcal{D}\mathcal{L}}{\max_j |D_j\mathcal{L}|}\right)\tau, \ \tau \in (0, \infty). \tag{4.4}$$

When $\tau = C\varepsilon^2 \max_j |D_j\mathcal{L}|$ the mapping gives $C\varepsilon^2\mathcal{D}\mathcal{L}$, it has the property that it parameterises the mapping $\tau \mapsto \mathcal{D}\mathcal{L}\tau$ over the co-ordinate direction where D_j has the component of largest magnitude. Our aim is now to describe the shift of the current node in terms of multiples of δ along this co-ordinate direction. First we define the quantity $G(\delta) := \text{Ceil}(C\varepsilon^2 \max_j |D_j\mathcal{L}|/\delta)$, this describes how many δ-steps in the mapping 4.4 it would take to achieve a shift by approximately $C\varepsilon^2\mathcal{D}\mathcal{L}$; the ceiling function in the

definition of $G(\delta)$ means that it may no longer recover $C\varepsilon^2\mathcal{DL}$ exactly. Finally we define the mapping

$$\Sigma(\xi) := A(\mathcal{N})\delta \left[\frac{\xi}{\delta}\right] \tag{4.5}$$

which snaps the vector ξ to the nearest point in $\delta\mathbb{Z}^d$ and then changes the co-ordinates from the frame along the curve to the standard basis in $\mathbb{R}^d$. Set $\mathcal{S} := \Sigma \circ S$. We now determine the optimal position for the currently selected node by testing the length for $q_i + \mathcal{S}(M_1 G(\delta)\delta)$, and finding M_1 such that the length of the curve in the updated position is reduced. Once $M_1 \in \mathbb{N}_0$ is found we then determine an $M_2 \in \mathbb{N}_0$ such that $q_i + \mathcal{S}(M_1 G(\delta)\delta + M_2\delta)$ also reduces length. The search for M_2 is performed using a binary search between 0 and the nearest number of the form 2^k to $G(\delta)$; the reason to choose these limits for the search is to ensure the binary search runs effectively with minimal rounding operations. Both M_1 and M_2 can be attained in a finite number of steps by [SZb, Proposition 3.5]. Should $M_1 = M_2 = 0$ then advancing in the gradient direction does not reduce length and one updates the node position with the single δ-step determined in the calculation of $\mathcal{DL}$, otherwise set $q_i + \mathcal{S}(M_1 G(\delta)\delta + M_2\delta) \to q_i$. Finally one proceeds to the next node or terminates the procedure if no further length reductions are possible.

By ensuring that the new node updates remain on the grid prescribed by the step methods we retain the consistency result of the step method. However, rather than determining how many δ steps one should take by a linear search, we use a predictor to reduce the number of new node positions we have to test. The reduction in the number of calculations is made in two ways, firstly, we are not restricted to moving in one basis direction per iteration. Using the gradient we are now able to move in multiple basis directions at a time, which potentially cuts the number of iterations needed to shorten the curve for high dimensional problems. The second is that the gradient approximation provides a faster estimate for how far to move the node for each test, reducing the number of δ-steps tested. This potentially reduces the number of tests needed by using the multiple step method. The performance of this procedure is measured against the other local methods in section 4.5.

Discussion of the Pseudocode for the Predictor-Corrector Procedure

The process initiates with procedure 4.9 which handles the order in which nodes are selected for movement. The procedure also determines a basis for the space of normal directions to the line joining q_a to q_b in an identical fashion to the step methods. Procedure 4.9 also handles the stopping condition, which as with the step methods is to stop once no node can be moved to further reduce length. Once a node is selected, in line 8,

the procedure continues to 4.10 to attempt to reduce length.

Procedure 4.9: Local Node Handling for Hybrid Curve Shortening Procedure

Inputs : Node $\in \mathcal{M}^{N\times d}(\mathbb{R})$: Each column representing a d-dimensional vector corresponding to a position along the initial curve. It will always hold that Node$(1) = q_a$ and Node$(N) = q_b$.
$\delta \in (0, 1)$: Selected as described in the introductory text.
$C \in (0, \infty)$: Selected as described in the introductory text.

Output: Node $\in \mathcal{M}^{N\times d}(\mathbb{R})$: Each column representing a d-dimensional vector corresponding to a position along the calculated curve. It will always hold that Node$(1) = q_a$ and Node$(N) = q_b$.

```
1  NodeOrder ← {3, 5, ..., 2N − 1} ∪ {2, 4, ..., 2N};
2  Tangent ← q_b − q_a;
3  {e_1, ..., e_{d−1}} ← FindNormalBasis(Tangent) ;
4  NewMin ← 1;
5  while NewMin ≠ 0 do
6  |  NewMin ← 0;
7  |  for i ← 1 to N − 2 do
8  |  |  ⟨ Node Movement — Procedure: 4.10 ⟩;
9  |  end
10 end
11 ReturnMidpoint(Node);
```

Procedure 4.10 is dedicated to the calculation of the optimal direction in which to move the selected node to reduce length. It starts by initialising the necessary variables for the calculations in line 3; in particular *Steps* records the number of δ-steps the node can move to reduce length. The current length is computed in line 4. In order to compute the gradient $\mathcal{DL}$ we use the arrays $FLen$ and $BLen$ to store the values of the L^+ and L^- respectively. The procedure proceeds to compute L^+ and L^- for each basis direction. The mechanism to determine if there exists a single step in a co-ordinate direction that reduces length is handled in line 9. If such a single step fails to exist then the procedure moves onto the next node, otherwise we proceed into the selection statement at line 13. If a single δ-step could reduce length, then we will determine if there is any additional benefit to moving in the gradient direction using procedure 4.11, otherwise the process moves to the next node. The flag $NewMin$ is set to 1 if there exists a single δ-step, otherwise it remains 0 as set in procedure 4.9.

Procedure 4.11 handles computing the gradient $\mathcal{DL}$ and re-scaling it as described in the mathematical discussion. The calculation of the rescaled gradient is performed in line 3. The function `GreatestDescent` in line 6 determines which component of *Direction* has the greatest magnitude and returns the component index and sign. The variable *Direction* is then rescaled again to behave like a graph over the largest component.

Procedure 4.10: Procedure to Determine the Direction that Minimises Length for the Hybrid Procedure.

```
1  ⟨ Called in line 8, Procedure: 4.9 ⟩;
2  ⟨ Input: ε, δ, C, NodeOrder, Node, NewMin ⟩;
3  MinJ ← 0, I ← NodeOrder(i), Steps ← 0;
4  BLen ← Len(Node(I-1),Node(I),Node(I+1));
5  FLen(1) ← 0, RLen(1) ← 0;
6  for j ← 1 to d − 1 do
7  |  FLen(j+1) ← Len(Node(I-1),Node(I)+δe_j,Node(I+1));
8  |  RLen(j+1) ← Len(Node(I-1),Node(I)−δe_j,Node(I+1));
9  |  if FLen < BLen ∨ RLen < BLen then
10 |  |  NewMin ← 1;
11 |  end
12 end
13 if NewMin = 1 then
14 |  ⟨ Compute Rescaled Gradient — Procedure: 4.11 ⟩;
15 |  ⟨ Test Node Positions — Procedure: 4.12 ⟩;
16 end
17 ⟨ Return to line 8, Procedure: 4.9 ⟩;
```

With this done the procedure proceeds to test how far the node needs to move in the gradient direction to reduce length in procedure 4.12.

Procedure 4.11: Procedure to compute the rescaled gradient.

```
1  ⟨ Called in line 14, Procedure: 4.10 ⟩;
2  ⟨ Input: ε, δ, C, FLen, RLen ⟩;
3  Direction ← −(Cε/2δ) · (FLen − RLen);
4  { MinJ, sMinJ } ← GreatestDescent(Direction);
5  GradientSteps ← Ceil(Abs(Direction(MinJ))/δ);
6  Direction ← sMinJ · Direction/Direction(MinJ);
7  ⟨ Return to line 14, Procedure: 4.10 ⟩;
```

Procedure 4.12 is a simple procedure to determine how many gradient lengths to move. The flag $LenMin$ determines when to stop moving the node in multiples of the gradient direction. The while loop in line 4 controls how many gradient lengths to first advance, the testing for optimality being handled by procedure 4.13. After this, should the number of δ-steps along the largest co-ordinate direction be greater than one, we can determine how many additional δ-steps we can make in the gradient direction, this is controlled by line 7 and procedure 4.14.

Procedure 4.13 controls the movement of the node along gradient lengths. The

vector by which the node should be shifted is constructed from the variables *Direction* and *Steps* using the function `ProduceShiftVector`. The function computes the vector in the direction of *Direction* and the magnitude of the longest component is $Steps \times \delta$, it also rotates the resulting vector into the standard basis of $\mathbb{R}^d$, rather than the frame along the initial curve. The remainder of procedure 4.13 is dedicated to a simple linear search to determine how many multiples of the gradient length we may move the node by to reduce length.

Should the procedure need to search within a gradient length for the optimal node position, then it uses procedure 4.14. Procedure 4.14 is a binary search procedure of approximately a gradient length. The range of the binary search is between 0 and `CeilBase2`(*GradientSteps*). The function `CeilBase2` takes an integer as an argument and returns the smallest integer of the form 2^k that is greater than or equal to the input. The additional condition in the while statement of line 6 ensures that the search stops when $MaxGuess = 2$ and $MinGuess = 1$.

Finally once, the number of δ-steps to be made in the gradient direction is determined to remains to move the node into the optimal position, using procedure 4.15. If it turns out that no step in the gradient direction will reduce length, then move in the δ-step that does reduce length. Otherwise, we reconstruct the shift vector that reduces length and shift the node by this. The process then returns to procedure 4.10 where either the next node is selected or the process is terminated.

Procedure 4.12: Procedure to Determine the Distance to Travel in the Optimal Direction

```
1  ⟨ Called in line 15, Procedure: 4.10 ⟩;
2  ⟨ Input: GradientSteps ⟩;
3  LenMin ← 1;
4  while LenMin = 1 do
5  |  ⟨ Test in Multiples of Gradient Distance — Procedure: 4.13 ⟩;
6  end
7  if GradientSteps > 1 then
8  |  ⟨ Bisect a Gradient Distance for Further Reduction — Procedure: 4.14 ⟩;
9  end
10 ⟨ Return to line 15, Procedure: 4.10 ⟩;
```

Procedure 4.13: Advance Current Node in Multiples of the Gradient Length

```
1  ⟨ Called in line 5, Procedure: 4.12 ⟩;
2  ⟨ Input: Direction, I, Node, BLen ⟩;
3  LenMin ← 0;
4  Steps ← Steps + GradientSteps;
5  TShift ← ProduceShiftVector(Direction, Steps);
6  TLen ← Len(Node(I-1),Node(I)+TShift,Node(I+1));
7  if TLen < BLen then
8  |  BLen ← TLen, LenMin ← 1;
9  else
10 |  Steps ← Steps − GradientSteps;
11 end
12 ⟨ Return to line 5, Procedure: 4.12 ⟩;
```

Procedure 4.14: Bisect a Single Gradient Length

```
1  ⟨ Called in line 8, Procedure: 4.12 ⟩;
2  ⟨ Input: Direction, I, Node, BLen, GradientSteps, Steps, LenMin ⟩;
3  MaxGuess ← CeilBase2(GradientSteps);
4  MinGuess ← 0;
5  MidGuess ← MaxGuess/2;
6  while MaxGuess > MinGuess ∧ MidGuess = Round(MidGuess) do
7  |  TShift ← ProduceShiftVector(Direction,MidGuess + Steps);
8  |  TLen ← Len(Node(I-1),Node(I)+TShift,Node(I+1));
9  |  if TLen < BLen then
10 |  |  BLen ← TLen, LenMin ← 1;
11 |  |  MinGuess ← MidGuess;
12 |  |  MidGuess ← (MaxGuess + MinGuess)/2;
13 |  else
14 |  |  MaxGuess ← MidGuess;
15 |  |  MidGuess ← (MaxGuess + MinGuess)/2;
16 |  end
17 end
18 Steps ← Steps + MinGuess;
19 ⟨ Return to line 8, Procedure: 4.12 ⟩;
```

Procedure 4.15: Procedure to Move Node to Optimal Length Reducing Position

```
1 ⟨ Called in line 10, Procedure: 4.12 ⟩;
2 ⟨ Input: MinJ, sMinJ, Direction, I, Node, Steps, TShift ⟩;
3 if Steps = 0 then
4 |  Node(I) ← Node(I)+sMinJ · δ · e_MinJ;
5 else
6 |  TShift ← ProduceShiftVector(Direction, Steps);
7 |  Node(I) ← Node(I)+TShift;
8 end
9 ⟨ Return to line 15, Procedure: 4.10 ⟩;
```

4.5 Comparison of the Local Procedures

In this section we study a specific family of examples for which the solutions are known. The specific example we will study is to calculate geodesics for the length functional

$$L_\alpha(u) = \int_0^1 e^{-\alpha \hat{n}\cdot u(\tau)} \|u'(\tau)\| d\tau,$$

where $n := \{0, 1, 1, \dots, 1\} \in \mathbb{R}^d$ and $\hat{n} := n/\|n\|$. In particular we will compute the geodesic joining $\{-1, 0, 0, \dots, 0\} \in \mathbb{R}^d$ to $\{1, 0, 0, \dots, 0\} \in \mathbb{R}^d$. Set

$$f_i(x) = \frac{1}{\alpha\sqrt{d-1}} \log\left(\frac{\cos(\alpha x)}{\cos(\alpha)}\right), \quad i = 1, \dots, d-1,$$

then $(x, f_1(x), \dots, f_{d-1}(x))$ describes the geodesic as a graph. Here we have two parameters that we can adjust to gain insight into the procedures. Most importantly we can vary the dimension of the problem and therefore we are able to determine the accuracy and efficiency of the procedures as the dimension of the problem increases.

The implementation of the procedures was performed in Matlab version 7.9.0.529. The hardware used to run the procedure was a Macintosh 'iMac' operating under the Mac operating system version 10.7.4. The processor is an Intel Core i5, that is four cores operating at 2.5GHz. Due to the relatively small amount of operating memory required the only other component that could effect performance is the RAM, which is 2 $\times$ 2GB 1333MHz DDR3. Other hardware factors we expect will have negligible impact on the performance of the procedure. We will study the single step, multiple step and the predictor-corrector methods in terms of performance and accuracy. We do not compare these methods with the gradient descent approach as discussed above.

At first, we will study all three methods, with the underlying dimension fixed at 3.

We will study the relationship between the number of nodes, that we will set as $2^N + 1$, against both the computed error to the actual solution and the time it takes to determine this solution. Our numerical computations consider $N \in \{4, 5, 6, 7, 8, 9\}$, for $N = 10$ we expect the duration of the calculation to be of the order of about a week.

We provide two plots that indicate performance. The first is related to a plot of the number of nodes against the average error between the calculated curve vertices and the true geodesic. As the solution is given as a graph over the first co-ordinate axis it is relatively easy to compute the desired error

$$E_N = \frac{1}{N} \sum_{i=0}^{N} \|q_i - (q_{i,1}, f(q_{i,1}))\|$$

where $q_{i,1}$ is the first component of the node q_i. We then plot $\log_2(2^N + 1)$ against $\log_2(E_N)$. The second plot is $\log_2(2^N + 1)$ against $\log_2(T_N)$ where T_N is the execution time for $2^N + 1$ nodes in seconds. Every time the program was running care was taken to ensure that only the procedure was running on the system at the time. When performing these computations we set $\alpha = 0.6$ and the problem dependent constant at $C = 4$.

4.5.1 Analysis for a 3 Dimensional Problem

Let us first analyse the error data. Both the multiple step and predictor-corrector methods produce similar calculated curves as the single step method. The next step is to analyse the behaviour of the data to predict the asymptotic properties of the procedures. In the plot a linear least squares regression curve is plotted, the equation of the line is $-0.56x - 1.74$ with a coefficient of correlation of $R^2 = 0.97$. The linear regression suggests that $E_N \sim CN^{-1/2}$ as $N \to \infty$.

Now let us analyse the execution time. The single step method has a linear least squares regression of $4.00x - 19.46$ with a coefficient of correlation of $R^2 = 0.97$, consequently the data is well modelled by the line. Should the trend continue then this suggests an asymptotic behaviour of $T_N \sim C_1 N^4$. The last two data points suggest however that the line may fall parallel with the other methods asymptotically. The key factor between the performance of these procedures is the constant that defines the lines.

The multiple step method has a linear least squares regression line of $3.57x - 17.56$ with a coefficient of correlation of $R^2 = 1.00$ to two decimal places, therefore the linear model is an excellent description of the data. This trend suggests that the procedure has an asymptotic behaviour of $T_N \sim C_2 N^{11/3}$.

The predictor-corrector method has a linear least squares regression line of $3.57x - 17.00$ with a coefficient of correlation of $R^2 = 1.00$ to two decimal places, therefore

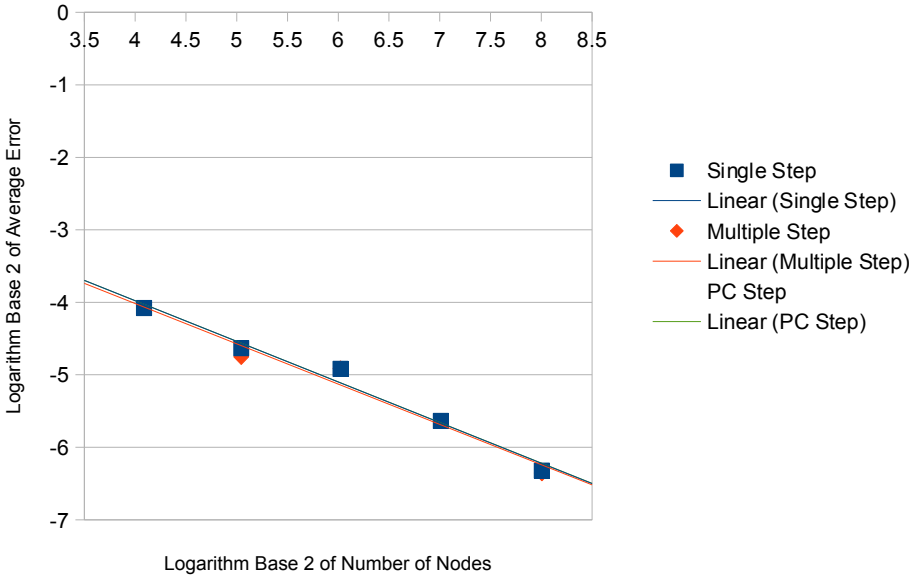

Figure 4-2: *Average Error for the 3 Dimensional Problem*

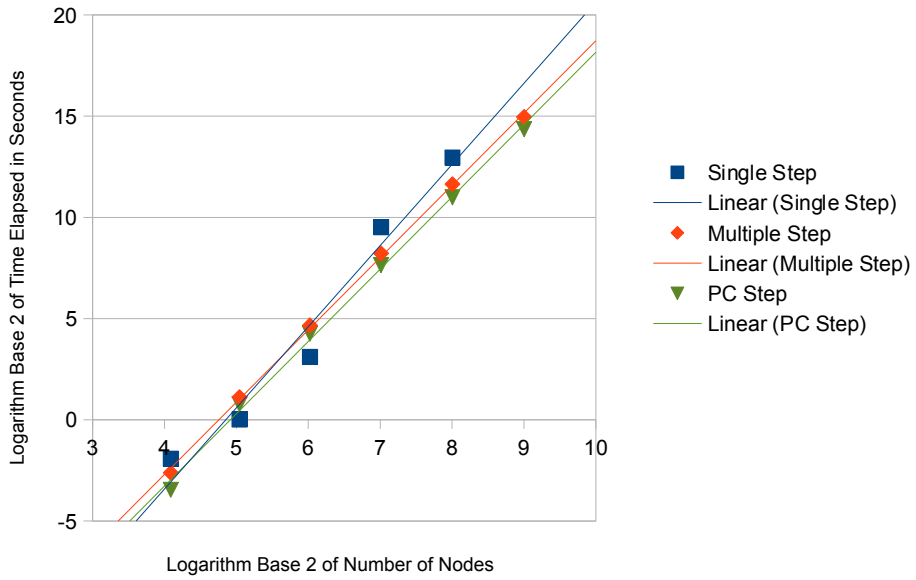

Figure 4-3: *Time Elapsed for the 3 Dimensional Problem*

the linear model is an excellent description of the data. This trend suggests that the procedure has an asymptotic behaviour of $T_N \sim C_3 N^{11/3}$ identical to the multistep procedure.

The only significant difference between the multiple step and predictor-corrector methods is the value of the constants C_2 and C_3. Whilst the asymptotic growth between these methods is predicted to be similar, the value of the constants has the effect of a 33% speed-up in terms of real time.

4.5.2 Analysis for Higher Dimensional Problems

It is clear from the analysis of the three procedures that the multiple step method and the proposed predictor-corrector method exhibit the best performance, for the remainder of this section we will only compare these two. For the three dimensional example we have seen that the multistep procedure and the predictor-corrector method exhibit similar performance. The next part of this analysis will determine whether increasing dimension also has an effect on the asymptotics. This is a reasonable investigation since a key difference between the two procedures is that the predictor-corrector method can move diagonally on the grid, whereas the multiple step method cannot.

In 3 dimensions, the geodesic is furthest from the initial curve when the first component is 0. When increasing the dimension of the problem, without changing the value for α, the geodesic becomes closer to the initial curve. For a given dimension we choose α to ensure that the furthest point between the geodesic and the initial curve is constant. Specifically, we choose α such that

$$\frac{1}{\alpha\sqrt{d-1}} \log\left(\frac{\cos(\alpha x)}{\cos(\alpha)}\right) = \frac{1}{\alpha\sqrt{2}} \log\left(\frac{1}{\cos(\alpha)}\right).$$

As a result of choosing α in this way we can ensure that the amount of work generated by adding additional dimensions is uniform. This also ensures that we are not considering the case of a lower dimensional problem embedded into a higher dimensional metric, where the calculations may be significantly easier.

Our additional higher dimensional calculations are performed in dimensions 5 and 9. We first discuss the performance for each dimension respectively and then discuss the behaviour of the procedures as the dimension changes. In this subsection all numbers are given to two decimal places. As with the 3 dimensional problem, when performing these computations we set $\alpha = 0.6$ and the problem dependent constant at $C = 4$.

Dimension 5

The error for both methods is identical. The least squares regression line for the error data takes the form $-0.47x - 3.61$ with a coefficient of correlation of $R^2 = 0.96$. It follows, as with the 3 dimensional case that $E_N \sim CN^{-1/2}$ as $N \to \infty$. As for the execution time, the multiple step method has a linear least squares regression line of $3.53x - 4.38$ with a coefficient of correlation of $R^2 = 1.00$. This trend suggests that $T_N \sim C_2 N^{11/3}$. The predictor-corrector method has a linear least squares regression line of $3.54x - 5.97$ with a coefficient of correlation of $R^2 = 1.00$. This trend suggests that $T_N \sim C_3 N^{11/3}$ for both procedures.

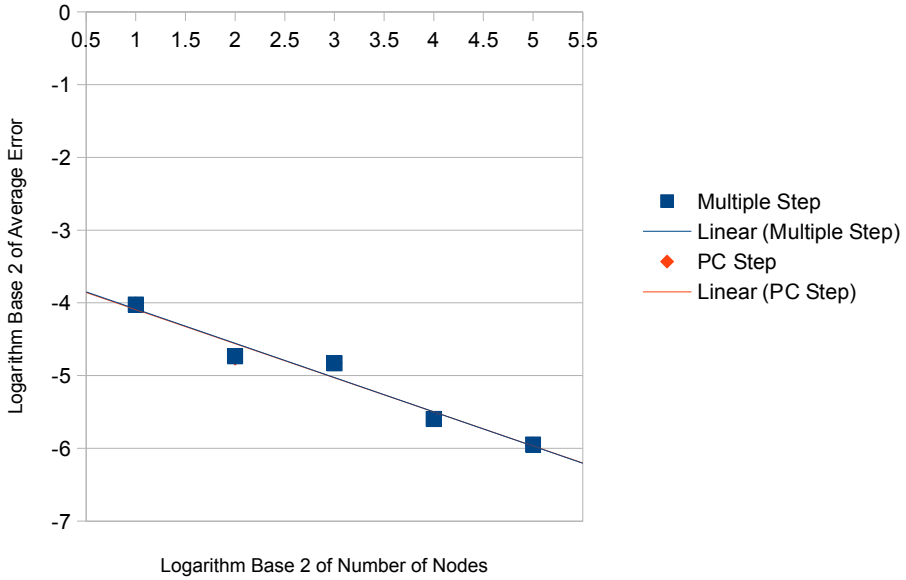

Figure 4-4: *Average Error for the 5 Dimensional Problem*

Dimension 9

The error for both methods is close but not identical in this case. This can arise as there can exist many local minimisers on the grid defined by the step methods. The least squares regression line for the error data of the multiple step method takes the form $-0.36x - 3.71$ with a coefficient of correlation of $R^2 = 0.85$. The least squares regression line for the error data of the predictor-corrector method takes the form $-0.36x - 3.63$ with a coefficient of correlation of $R^2 = 0.89$. As can be seen the least squares regression lines are approximately parallel; therefore for either procedure $E_N \sim CN^{-1/3}$ as $N \to \infty$. We now analyse the execution time. The multiple step method has a linear least squares

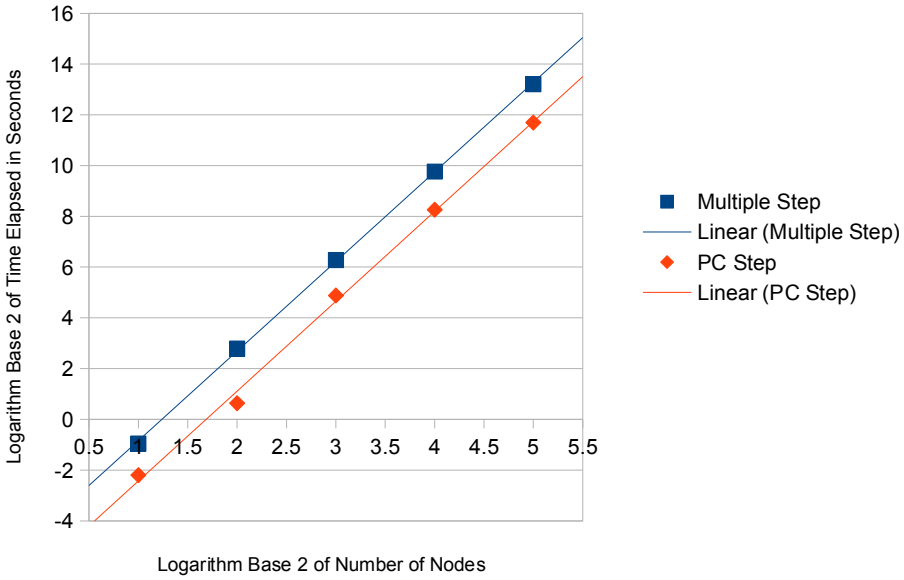

Figure 4-5: *Time Elapsed for the 5 Dimensional Problem*

regression line of $3.58x - 13.52$ with a coefficient of correlation of $R^2 = 1.00$. This trend suggests that $T_N \sim C_2 N^{11/3}$. The predictor-corrector method has a linear least squares regression line of $3.42x - 15.13$ with a coefficient of correlation of $R^2 = 1.00$. This trend, as before suggests that $T_N \sim C_3 N^{11/3}$ for both procedures.

4.5.3 Relative Performance in Dimension

As can be seen from examining figures 4-3, 4-5 and 4-7, whilst the asymptotics of the procedures are very similar, the distance between these lines seemingly increases with dimension. The exact separation between regression lines is 0.56,1.59 and 1.61 in 3,5 and 9 dimensions respectively. While there is little additional benefit between 5 and 9 dimensions, it is clear that there is an added benefit to using the predictor-corrector method. In the remaining sections we will make our computations using only the predictor-corrector method.

4.6 Parallelisation of the Global Procedure

The Birkhoff procedure is a process that is natural to parallelise. We have the global procedure, that handles the positions of the nodes that describe the extended curve. Whereas we also have the local procedure that determines an approximation of the local geodesic joining the nodes and returns a new node position. In order to gain a significant

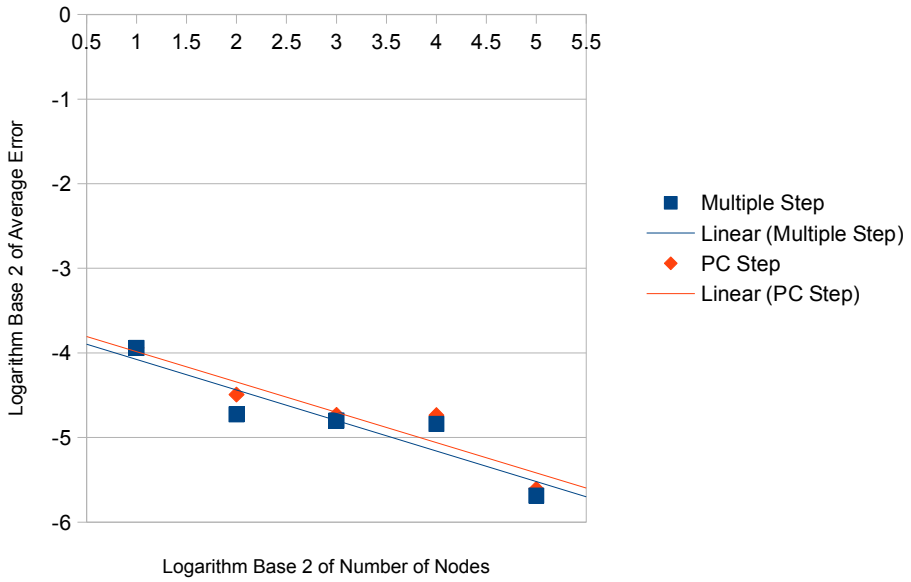

Figure 4-6: *Average Error for the 9 Dimensional Problem*

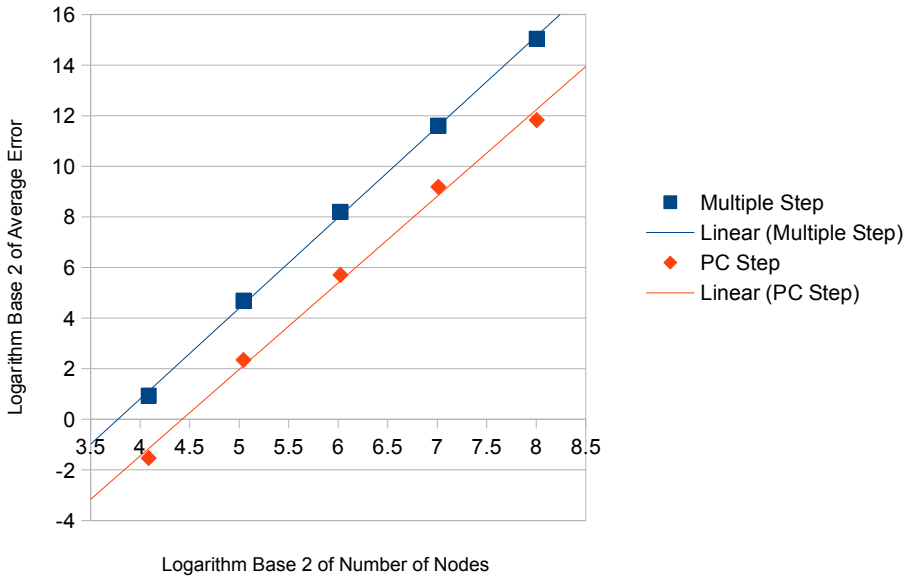

Figure 4-7: *Time Elapsed for the 9 Dimensional Problem*

speed up of the procedure we can distribute the computational workload across several processing cores. While there may be several ways to implement a parallel Birkhoff procedure we focus on one.

Discussion of the Pseudocode for the Parallel Global Procedure

We will designate one core to the processing of the global procedure. The remaining cores will be dedicated to computing the local procedure and returning new node positions to the core running the global procedure. One issue that may arise when running the Birkhoff procedure is that, due to the fact that the movement of a node depends on it's neighbouring nodes, one may encounter the situation where a core needs to know a node position which is being updated by another node. To prevent this we employ the rudimentary measure that all of the cores running the local procedure first move all of the even numbered nodes, once each of these nodes have been tested to reduce length, then the cores move the odd numbered nodes. The process is described in procedure 4.16 and 4.17, it assumes that we are using a system with $C_{\max} \geq 2$ cores.

Analysis of the Parallel Procedure

The implementation of the procedures that we measure here was performed in C with the MPI package for the parallel features. The hardware used to run the procedure was as described in the analysis of the step procedures.

For the comparison of the methods we measure the time it takes to execute the procedure. We will compute the solutions to the example problem as in section 4.5. We study higher dimensional problems, namely the example in 12 and 24 dimensions, to demonstrate the performance of the procedures in these settings. The parameters for the local procedure were chosen to be $\alpha = 0.6$ and $C = 4$. The calculation was made using 9 global nodes, to reduce the error we increased the number of local nodes for each calculation. The error and time taken were plotted against the total number of nodes, that is, the global nodes with any additional local nodes between them.

Let us first discuss the error data for both the 12 and 24 dimensional problems. For the purposes of determining the asymptotic behaviour we ignore the outlying data point. The least squares regression line for the error data of the 24 dimensional problem takes the form $-0.20x - 3.43$ with a coefficient of correlation of $R^2 = 0.98$. This suggests that $E_N \sim CN^{-1/5}$ as $N \to \infty$. The least squares regression line for the error data of the 12 dimensional problem takes the form $-0.10x - 4.73$ with a coefficient of correlation of $R^2 = 0.96$. This suggests that $E_N \sim CN^{-1/10}$ as $N \to \infty$. It would prove to be an interesting challenge to attempt to derive rigourous error bounds for the procedures we

Procedure 4.16: Parallel Global Curve Shortening Procedure (Core 1)

Inputs : CurveData $\in \mathcal{M}^{(2N+1)\times d}(\mathbb{R})$: Each row representing a d-dimensional vector corresponding to a position along the initial curve.
TolMin $\in (0,1)$: Sets stopping condition for the procedure.
Output: CurveData $\in \mathcal{M}^{(2N+1)\times d}(\mathbb{R})$: Each row representing a d-dimensional vector corresponding to a position along the calculated curve.

```
Tol ← 1;
NNumber ← {3, 5, ..., 2N − 1} ∪ {2, 4, ..., 2N};
while Tol > TolMin do
    for i ← 1 to N − 1 do
        I = NNumber(i);
        ⟨ Send CurveData(I-1) and CurveData(I+1) to any available core, Procedure: 4.17 ⟩;
        ⟨ Receive NPosition from any core, Procedure: 4.16 ⟩;
        Tol = Tol + Norm (NPosition − CurveData(I));
        CurveData(I) = NPosition;
    end
    for i ← N to 2N − 1 do
        I = NNumber(i);
        ⟨ Send CurveData(I-1) and CurveData(I+1) to any available core, Procedure: 4.17 ⟩;
        ⟨ Receive NPosition from any core, Procedure: 4.16 ⟩;
        Tol = Tol + Norm (NPosition − CurveData(I));
        CurveData(I) = NPosition;
    end
    Tol = Tol/N;
end
```

Procedure 4.17: Parallel Global Curve Shortening Procedure (Core $2, \ldots, C_{\max}$)

```
⟨ Receive CurveData(I-1) and CurveData(I+1) from Core 1, Procedure: 4.16 ⟩;
NPosition ← LocalProcedure (CurveData(I-1), CurveData(I+1),Params);
⟨ Send NPosition to Core 1, Procedure: 4.16 ⟩;
```

describe here.

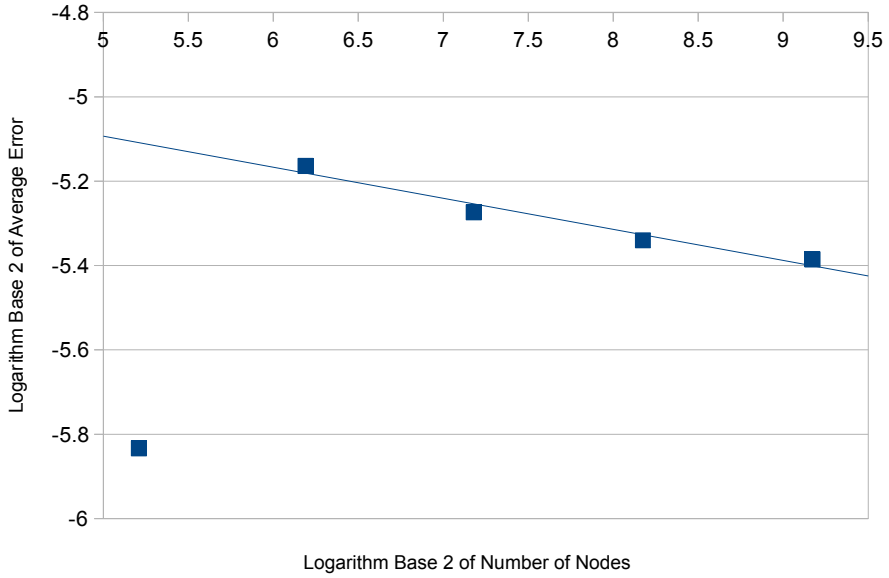

Figure 4-8: *Average Error for the 12 Dimensional Problem in the Parallel Implementation*

It remains to analyse the execution time for the 12 and 24 dimensional problems. The least squares regression lines for the execution time of the 24 dimensional problem take the forms $3.56x - 3.44$, $3.52x - 4.06$ and $3.61x - 4.49$ for the case of 2, 3 and 4 cores respectively. The coefficient of correlation is $R^2 = 1$ in all cases. This suggests that $T_N \sim CN^{11/3}$ as $N \to \infty$. The least squares regression lines for the execution time of the 12 dimensional problem take the forms $3.48x - 5.20$, $3.44x - 5.81$ and $3.52x - 6.27$ for the case of 2, 3 and 4 cores respectively. The coefficient of correlation is $R^2 = 1$ in all cases. This suggests that $T_N \sim CN^{7/2}$ as $N \to \infty$.

Let us now analyse the effect that increasing the number of cores has on the performance of the procedure. The case of 2 cores is effectively running the global procedure on one core, with the computation intensive local procedure running on the other core. Despite this, a comparison of figures 4-7 and 4-10 shows that they run at approximately the same speed. This arises due to the fact that the parallel implementation is made in C. It is immediately clear from figures 4-10 and 4-11 that there is a computation benefit in using 3 cores rather than three. The separation between the regression lines being 0.54 and 0.61 in the 12 and 24 dimensional problems respectively. This is due to the obvious reason that the calculation intensive local procedures can be shared over two processors. Unfortunately this effect doesn't seem to persist for the case of 4 cores. While there is some added benefit for a low total number of nodes, this effect disappears for additional

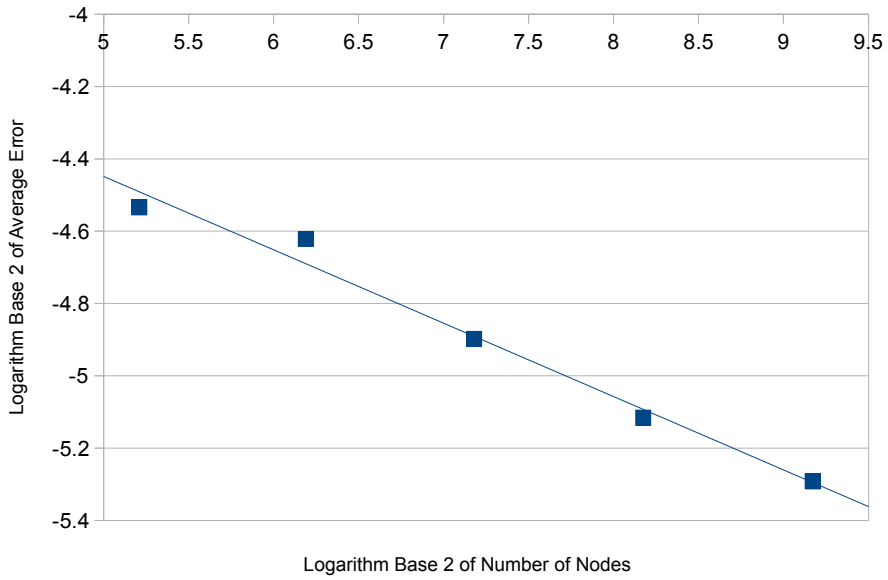

Figure 4-9: *Average Error for the 24 Dimensional Problem in the Parallel Implementation*

nodes. The reason for this is owing to the particular nature of the parallel procedure we have implemented. The procedure, to ensure that each node is moved according to the correct information about it's neighbouring nodes, first moves all the even nodes in parallel before moving all of the odd nodes in parallel. It is due to this method that a single core may hold up the other free cores from proceeding to the odd nodes. It is intended for future work to write a more effective parallel procedure enables free cores to move nodes that are allowed to be moved. In order to do this we will ensure that the core running the global procedure keeps track of which cores are moving which nodes and assign free nodes to free cores as appropriate.

4.7 Using the Maupertuis Principle to Solve a Molecular Dynamics Problems

4.7.1 A Mathematical Model for Butane

To provide an interesting example of a higher dimensional molecular dynamics calculation we consider the motion of a Butane molecule. We use the model as presented in [Fis97], here the problem has 12 degrees of freedom, making it a good problem to study using the predictor-corrector method. In this model Butane is represented as four masses, representing either the CH_3 or CH_2 groups, see figure 4-15. Let $\mu = 1.66 \cdot 10^{-24} g$ denote the atomic mass unit. The masses of the CH_2 and CH_3 groups are 14μ and 15μ

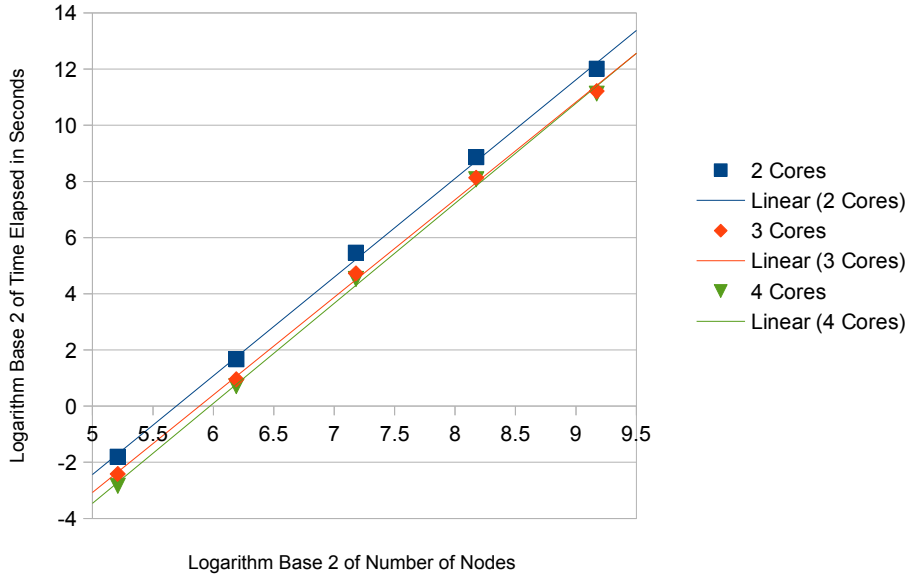

Figure 4-10: *Time Elapsed for the 12 Dimensional Problem in the Parallel Implementation*

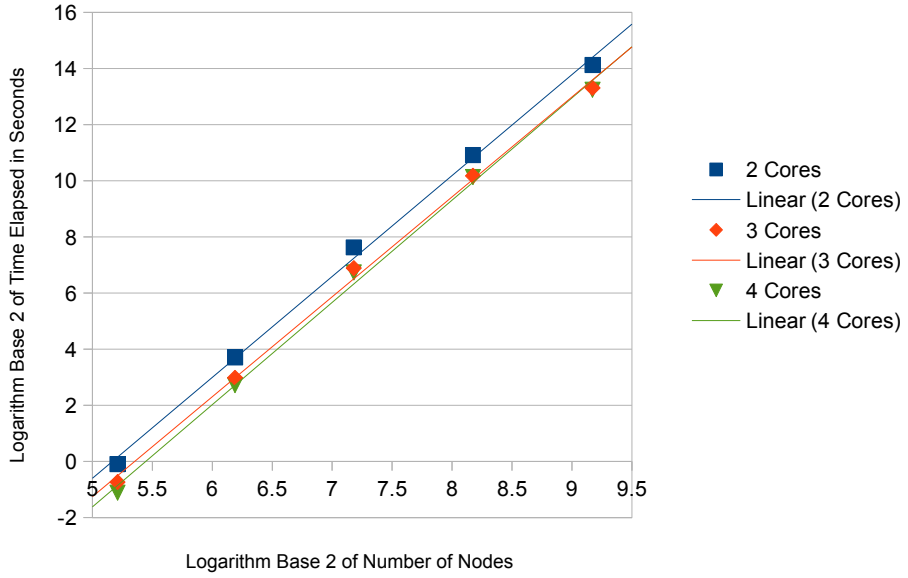

Figure 4-11: *Time Elapsed for the 24 Dimensional Problem in the Parallel Implementation*

respectively. Define the mass matrix of the system as

$$M := \begin{pmatrix} 15\mu I_3 & 0_{3\times 6} & 0_{3\times 3} \\ 0_{6\times 3} & 14\mu I_6 & 0_{6\times 3} \\ 0_{3\times 3} & 0_{3\times 6} & 15\mu I_3 \end{pmatrix}$$

where I_n is the $n \times n$ identity matrix and $0_{n\times m}$ is the $n \times m$ zero matrix. The potential function that describes the dynamics of the Butane molecule is the following

$$\begin{aligned} V_b(x) = \frac{\alpha}{2}\left(\left(\|q_1 - q_2\| - r_0\right)^2 + \left(\|q_2 - q_3\| - r_0\right)^2 + \left(\|q_3 - q_4\| - r_0\right)^2\right) \\ + \frac{\beta}{2}\left(\left(\cos\left(\frac{(q_2 - q_1)\cdot(q_3 - q_2)}{\|q_2 - q_1\|\|q_3 - q_2\|}\right) - \theta_0\right)^2 \right. \\ \left. + \left(\cos\left(\frac{(q_3 - q_2)\cdot(q_4 - q_3)}{\|q_3 - q_2\|\|q_4 - q_3\|}\right) - \theta_0\right)^2\right) \\ + v\sum_{i=1}^{6} a_i \cos\left(\frac{((q_2 - q_1)\wedge(q_3 - q_2))\cdot((q_3 - q_2)\wedge(q_4 - q_3))}{\|(q_2 - q_1)\wedge(q_3 - q_2)\|\|(q_3 - q_2)\wedge(q_4 - q_3)\|}\right)^{i-1}. \end{aligned}$$

The values for the parameters, taken directly from [Fis97], are $\alpha = 83.7$kcal, $r_0 = 1.53$Å, $\beta = 43.1$kcal, $v = 1.987$kcal and $\theta_0 = 109.5$. Furthermore, $a_1 = 1.116$, $a_2 = 1.462$, $a_3 = -1.578$, $a_4 = -0.368$, $a_5 = 3.156$ and $a_6 = -3.788$.

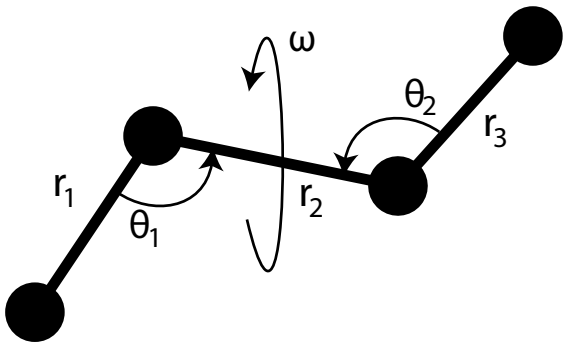

Figure 4-12: *An illustration of the mass-spring model for a Butane molecule, including the notation we use in this section. The outer solid balls represent CH_3 groups, while the inner balls represent CH_2 groups. The lines joining the balls represent harmonic springs describing $C - C$ bonds.*

4.7.2 A Simulation of the Change of Conformation in Butane

The aim of this section is to compute geodesics in the Jacobi metric

$$\int_0^1 \sqrt{2(E - V_b(u(\tau)))\langle u'(\tau), Mu'(\tau)\rangle}d\tau,$$

the appearance of the mass matrix M is due to the fact that the masses in the system are not identical [Arn97]. This does not affect the curve shortening procedure as $\sqrt{\langle \xi, M\xi\rangle}$ defines a norm on $\mathbb{R}^d$. The only modification necessary in so far as the implementation of the procedures is that the length approximation is now given by

$$L(q_a, q_b) := \left(\frac{a(q_a) + a(q_b)}{2}\right)\sqrt{\langle q_b - q_a, M(q_b - q_a)\rangle}, \tag{4.6}$$

where a denotes the metric coefficient.

Define the dihedral angle as

$$\omega = \frac{((q_2 - q_1) \wedge (q_3 - q_2)) \cdot ((q_3 - q_2) \wedge (q_4 - q_3))}{\|(q_2 - q_1) \wedge (q_3 - q_2)\|\|(q_3 - q_2) \wedge (q_4 - q_3)\|},$$

this describes the torsion about the central $C-C$ bond, it is this parameter that describes the change in conformation. In particular, we will compute the motion of the butane molecule as it moves between the gauche ($\omega = \pi/3$) and trans ($\omega = \pi$) configurations as defined in [Fis97]. Specifically, we start with the gauche configuration with

$$\begin{aligned} q_1 &= (-2.295, -1.530, 1.185),\\ q_2 &= (-0.765, -1.530, -1.186),\\ q_3 &= (0.000, 0.000, -1.186),\\ q_4 &= (1.530, 0.000, 1.185) \end{aligned}$$

and end with the trans configuration where

$$\begin{aligned} q_1 &= (-2.295, -1.530, 0.000),\\ q_2 &= (-0.765, -1.530, 0.000),\\ q_3 &= (0.000, 0.000, 0.000),\\ q_4 &= (1.530, 0.000, 0.000). \end{aligned}$$

The local procedure parameters were set to be $\alpha = 1$ and $C = 2$ for the following calculations.

In order to make the computation we are required to specify a value for E. Should

E be sufficiently large, then the geodesics become uniformly close to the straight line joining the initial and final states. Should E be too small, then there is the possibility that the geodesic connecting the gauche and trans states enters a region where the Jacobi metric degenerates. The subsequent implication for the physical trajectory is that it may no longer join the initial and final states. The search for an E between these extremes is performed by trial and error, using a smaller number of nodes. In the implementation of the procedure we have set a flag terminating the procedure should the geodesic be moved into a region where $E - V(x) < 0$, as this indicates that the energy is too low. To detect the values for E with geodesics close to the initial curve is performed as an observation on the time it took to compute the final curve. When the calculation took less than a minute to calculate, with 129 nodes, we considered the energy to be too high.

We run the procedure for a range of energies E. The results of these computations are illustrated in figure 4-13. To determine an appropriate energy we first performed the calculation with a relatively small number of both global and local nodes, 16 and 17 nodes respectively. Trajectories for energies 520kcal, 521kcal, 524kcal, 525kcal and 527kcal were calculated. The purpose of doing this was to determine for which energies the dynamics are relatively simple, that is, close to the initial curve, and those energies for which there may be no physical solution. We measure the difference between the initial curve and final geodesic by studying how ω changes along the curve. As can be seen in figure 4-13, the computations suggest that this transition between entering the degenerate region and being close to the initial curve is continuous in E. For energies in the region of 515kcal, the procedure still produces a solution with the smaller quantity of nodes, however, as the number of nodes increase the geodesic entered the boundary of the configuration manifold. For an energy of 520kcal we are able to compute, using a large number of nodes, a trajectory that doesn't enter the degenerate region of the metric. This solution is also sufficiently far away from the initial curve to justify making further calculations. To see the values of the metric coefficient for the larger calculation performed in the next section see figure 4-20.

4.7.3 Analysis of Simulation

Therefore we study the case where $E = 520$kcal, as the motion will potentially display some interesting dynamics, being sufficiently far away from the initial curve. Increasing the total number of nodes to 481 the results of the calculation can be seen in figure 4-14. Using this data we plot the motion along the geodesic in terms of the bond extensions and molecular angles. The calculated geodesic is taken to be parameterised according to normalised arc length. It took approximately an hour to run this simulation using the parallel C procedure on 4 cores. It took approximately 4 hours to run the local procedure

in MATLAB on a single core.

The bond extensions, denoted r_i is defined as $\|q_{i+1} - q_i\|$ for $i = 1, 2, 3$ are plotted in figures 4-15, 4-16 and 4-17. As can be seen, by comparing figures 4-15 and 4-17, the change in bond extensions r_1 and r_3 do deviate from the initial curve, and in fact they do so in an identical fashion, a symmetry worth noting. Another interesting feature of the change in bond length is the behaviour of r_2. The change in r_2 deviates significantly from the initial curve, and does so in a symmetric fashion. The bond angles θ_i are defined as

$$\cos(\theta_i) = \frac{\langle q_{i+1} - q_i, q_{i+2} - q_{i+1} \rangle}{\|q_{i+1} - q_i\| \|q_{i+2} - q_{i+1}\|}$$

for $i = 1, 2$. The bond angles are plotted in figures 4-18 and 4-19. The behaviour of the bond angle also significantly differs from the initial curve. However, again the way that the θ_i evolve is identical, another noteworthy symmetry. The angle begins initially at 75°, flexing wider to it's peak at approximately 85° before closing to the final angle of 65°.

From these calculations we gain an insight into how the butane molecule changes conformation, that it twists, as indicated by the change in ω, in a symmetric fashion. As it twists the bond angles increase before decreasing to their final state. The bond extensions r_1 and r_3 decrease identically, and monotonically towards their final states. Finally the bond extension r_2 increases to a maximum of approximately 1.9Å, at the same point where the θ_i attain their maximum, before decreasing to the final state.

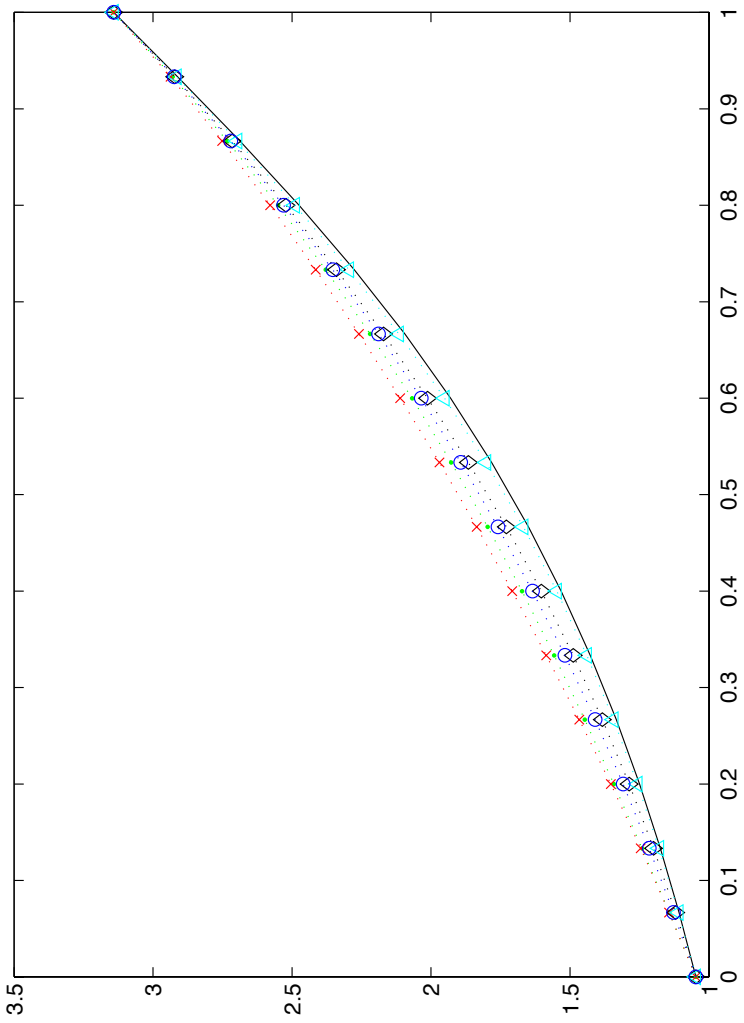

Figure 4-13: *A plot describing the change in ω (vertical axis) as a function of normalised arc length (horizontal axis) along the computed geodesic. The solid line represents the initial curve, the straight line joining the trans and gauche configurations. The remaining dotted curves represent the calculated geodesics with the markers indicating the final node positions for the global procedure. The corresponding energies for each computed trajectory, in order from the top curve to the bottom curve are, 520kcal, 521kcal, 524kcal, 525kcal and 527kcal.*

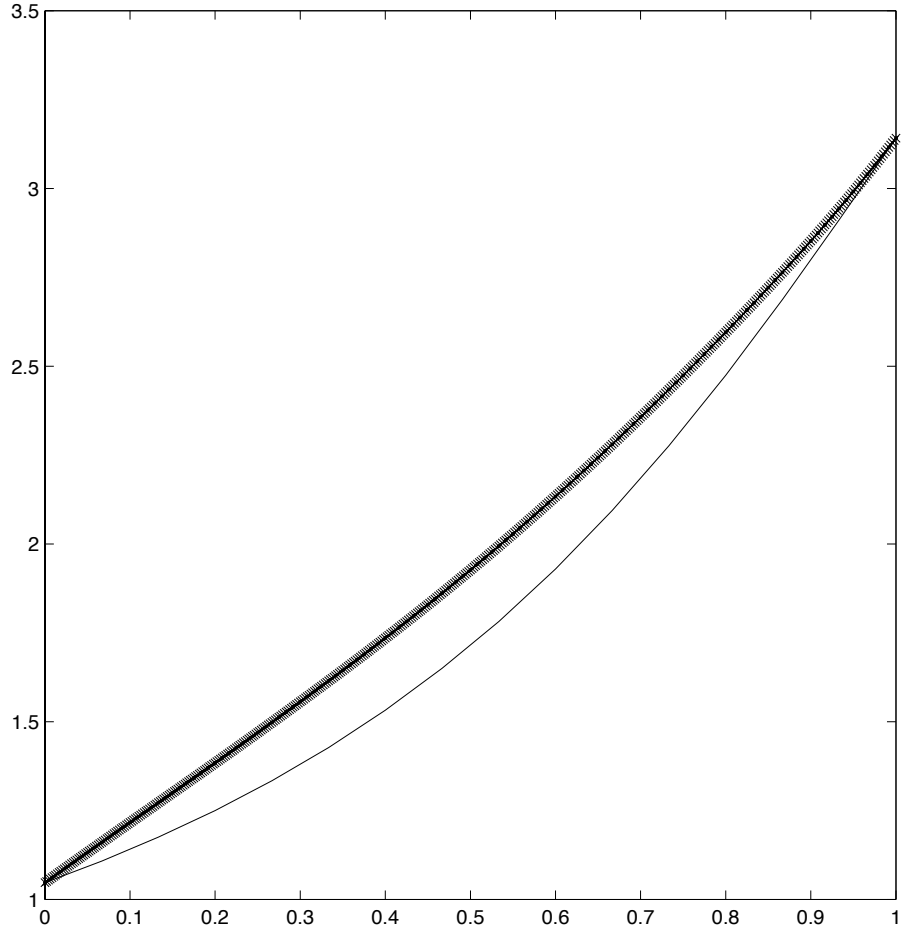

Figure 4-14: *A plot describing the change in ω for $E = 520kcal$ (vertical axis) as a function of normalised arc length (horizontal axis) along the computed geodesic. The calculation here is made with a total of 481 nodes. The solid line indicates the initial curve, the crosses indicate the computed geodesic.*

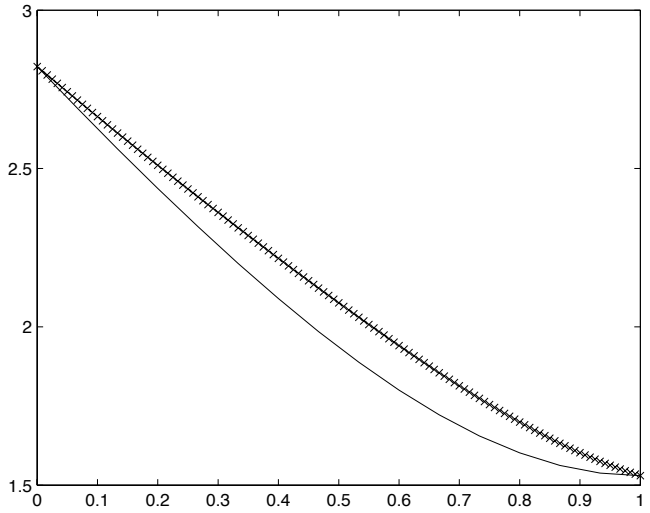

Figure 4-15: *A plot describing the change in r_1 (vertical axis) as a function of normalised arc length (horizontal axis) along the computed geodesic. The calculation here is made with a total of 481 nodes, of which 121 are shown. The solid line indicates the initial curve, the crosses indicate the computed geodesic.*

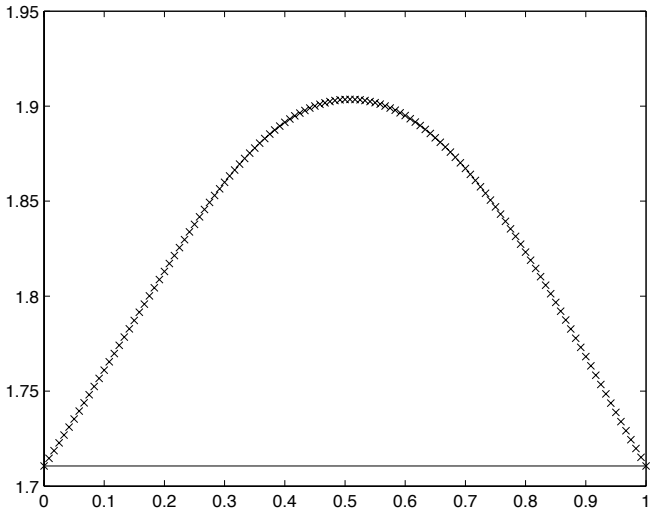

Figure 4-16: *A plot describing the change in r_2 (vertical axis) as a function of normalised arc length (horizontal axis) along the computed geodesic. The calculation here is made with a total of 481 nodes, of which 121 are shown. The solid line indicates the initial curve, the crosses indicate the computed geodesic.*

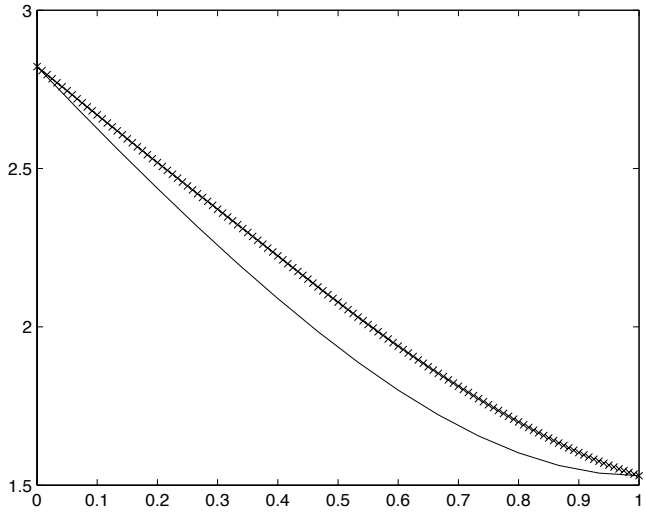

Figure 4-17: *A plot describing the change in r_3 (vertical axis) as a function of normalised arc length (horizontal axis) along the computed geodesic. The calculation here is made with a total of 481 nodes, of which 121 are shown. The solid line indicates the initial curve, the crosses indicate the computed geodesic.*

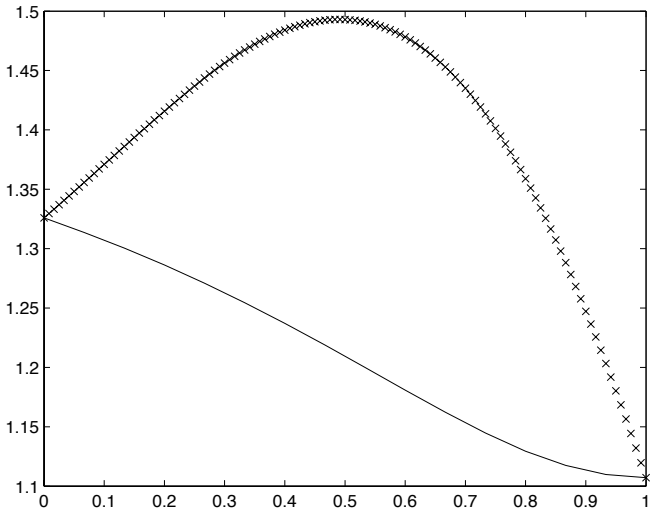

Figure 4-18: *A plot describing the change in θ_1 (vertical axis) as a function of normalised arc length (horizontal axis) along the computed geodesic. The calculation here is made with a total of 481 nodes, of which 121 are shown. The solid line indicates the initial curve, the crosses indicate the computed geodesic.*

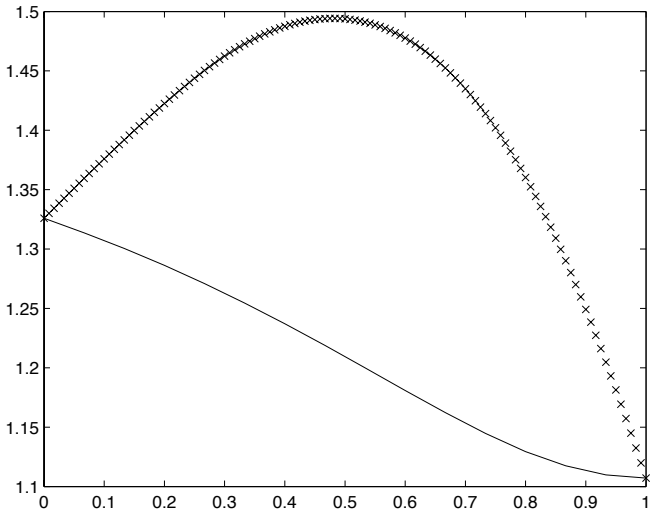

Figure 4-19: *A plot describing the change in θ_2 (vertical axis) as a function of normalised arc length (horizontal axis) along the computed geodesic. The calculation here is made with a total of 481 nodes, of which 121 are shown. The solid line indicates the initial curve, the crosses indicate the computed geodesic.*

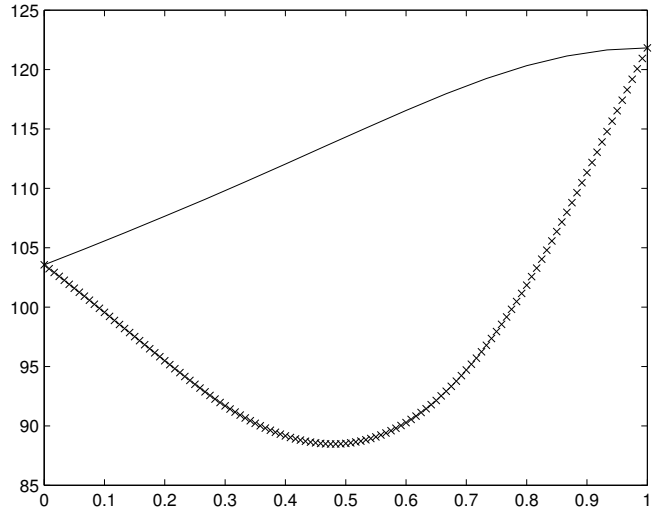

Figure 4-20: *A plot describing the change in $\sqrt{2(E - V_b(u(\tau)))}$ (vertical axis) as a function of normalised arc length (horizontal axis) along the computed geodesic. The calculation here is made with a total of 481 nodes, of which 121 are shown. The solid line indicates the initial curve, the crosses indicate the computed geodesic.*

Appendix A

MATLAB Code Listing of Curve Shortening Code

A.1 Local Algorithm Implementation

The purpose of this function is to perform the local algorithm. The parameter *NodeOrder* indicates that the calculation should be performed with $2^N + 1$ nodes, where N is *NodeOrder*. The variable *Results* returns the midpoint along the local geodesic. The variable *Alpha* takes values between 2.5 and 3. The other key variable, not seen in the pseudocode, is *MetricParameters*, which is essentially a free array which one may use to configure the metric function.

```
function [Curve_Data GraphFail Length] = GradientJacobi(NodeOrder,
    Start_Point,End_Point, Constant, Alpha, varargin)

Dimension = size(Start_Point,1);

if nargin - 5 == 0
    MetricParameters = 0;
else
    MetricParameters = varargin{1};
end

N = 2^NodeOrder+1;
N_1 = N-1;
Curve_Data = zeros(Dimension,N);
Local_Length = zeros(N_1,1);
ip1 = 2;
ip2 = 3;
Tangent_Step = (End_Point - Start_Point)/N_1;
```

```
Epsilon = norm(Tangent_Step);
Delta = Epsilon^Alpha;

Basis_Rotation_Matrix = FindFrame(Tangent_Step);
for i = 1:N
    Curve_Data(:,i) = Start_Point + (i-1) * Tangent_Step;
end

for i = 1:2:N-2
    [Local_Length(i) Local_Length(ip1)] =
        L(Curve_Data(:,i),Curve_Data(:,ip1),Curve_Data(:,ip2),
        Metric_Coefficient(MetricParameters,Curve_Data(:,i)),
        Metric_Coefficient(MetricParameters,Curve_Data(:,ip1)),
        Metric_Coefficient(MetricParameters,Curve_Data(:,ip2)));
    ip1 = ip1 + 2;
    ip2 = ip2 + 2;
end

Fwd_NewLength = zeros(Dimension,2);
Bwd_NewLength = zeros(Dimension,2);
Candidate_Shift = zeros(Dimension,1);
Advance_Direction = zeros(Dimension,1);
Shift_Gradient = zeros(Dimension,1);
Curve_Data_Moved = 1;

while (Curve_Data_Moved == 1)
    Curve_Data_Moved = 0;

    for i = [2:2:N,N-2:-2:3]
        im1 = i-1;
        ip1 = i+1;

        OldLength = Local_Length(im1) + Local_Length(i);
        Number_of_Steps = 0;
        Not_Local_Minimum = 0;

        for p = 2:Dimension
            [Fwd_NewLength(p,1) Fwd_NewLength(p,2)] =
                L(Curve_Data(:,im1),Curve_Data(:,i)+Delta
                *Basis_Rotation_Matrix(:,p),Curve_Data(:,ip1),
                Metric_Coefficient(MetricParameters,Curve_Data(:,im1)),
                Metric_Coefficient(MetricParameters,Curve_Data(:,i)+Delta
                *Basis_Rotation_Matrix(:,p)),
                Metric_Coefficient(MetricParameters,Curve_Data(:,ip1)));
            Forward_Length = Fwd_NewLength(p,1) + Fwd_NewLength(p,2);
```

```
    [Bwd_NewLength(p,1) Bwd_NewLength(p,2)] =
        L(Curve_Data(:,im1),Curve_Data(:,i)-Delta
        *Basis_Rotation_Matrix(:,p),Curve_Data(:,ip1),
        Metric_Coefficient(MetricParameters,Curve_Data(:,im1)),
        Metric_Coefficient(MetricParameters,Curve_Data(:,i)-Delta
        *Basis_Rotation_Matrix(:,p)),
        Metric_Coefficient(MetricParameters,Curve_Data(:,ip1)));

    Backward_Length = Bwd_NewLength(p,1) + Bwd_NewLength(p,2);

    if (Forward_Length < OldLength) | (Backward_Length < OldLength)
        Not_Local_Minimum = 1;
    end

    Advance_Direction(p) = -(1/(2 * Delta * Epsilon)) *
        (Forward_Length - Backward_Length);
end

if Not_Local_Minimum == 1
    Advance_Direction = Constant * Epsilon^2 * Advance_Direction;
    Long_Direction = 2;

    for j = 2:Dimension
        if (abs(Advance_Direction(j)) >
            abs(Advance_Direction(Long_Direction)))
            Long_Direction = j;
        end
    end

    TV = abs(Advance_Direction(Long_Direction))/Delta;
    Number_of_Gradient_Steps =
        ceil(abs(Advance_Direction(Long_Direction))/Delta);
    Single_Step_Direction =
        sign(Advance_Direction(Long_Direction));
    Shift_Gradient_Scaling =
        Single_Step_Direction/Advance_Direction(Long_Direction);
    for k = 2:Dimension
        Shift_Gradient(k) = Shift_Gradient_Scaling *
            Advance_Direction(k);
    end

    Length_Reduced = 1;

    while Length_Reduced == 1
        Length_Reduced = 0;
```

```
        Number_of_Steps = Number_of_Steps +
            Number_of_Gradient_Steps;
        Candidate_Shift(1) = 0;
        for k = 2:Dimension
            Candidate_Shift(k)= Delta * round(Shift_Gradient(k) *
                Number_of_Steps);
        end
        Candidate_Shift = Basis_Rotation_Matrix *
            Candidate_Shift;

        [NewLength_0 NewLength_1] =
            L(Curve_Data(:,im1),Curve_Data(:,i)+Candidate_Shift,
            Curve_Data(:,ip1),
            Metric_Coefficient(MetricParameters,Curve_Data(:,im1)),
            Metric_Coefficient(MetricParameters,Curve_Data(:,i)+
            Candidate_Shift),
            Metric_Coefficient(MetricParameters,Curve_Data(:,ip1)));
        NewLength = NewLength_0 + NewLength_1;

        if (NewLength < OldLength)
            Length_Reduced = 1;
            OldLength = NewLength;
            Local_Length(im1) = NewLength_0;
            Local_Length(i) = NewLength_1;
        else
            Number_of_Steps = Number_of_Steps -
                Number_of_Gradient_Steps;
        end
    end

    if Number_of_Gradient_Steps > 1
        Max_Guess = 2^(ceil(log2(Number_of_Gradient_Steps)));
        Min_Guess = 0;
        Mid_Guess = 0.5 * Max_Guess;

        while (Max_Guess > Min_Guess) & (Mid_Guess == round(Mid_Guess))
            Candidate_Shift(1) = 0;
            for k = 2:Dimension
                Candidate_Shift(k)= Delta * round(Shift_Gradient(k) *
                    (Mid_Guess+Number_of_Steps));
            end
            Candidate_Shift = Basis_Rotation_Matrix *
                Candidate_Shift;

            [NewLength_0 NewLength_1] =
                L(Curve_Data(:,im1),Curve_Data(:,i)+Candidate_Shift,
```

```
            Curve_Data(:,ip1),
            Metric_Coefficient(MetricParameters,Curve_Data(:,im1)),
            Metric_Coefficient(MetricParameters,Curve_Data(:,i)+
            Candidate_Shift),
            Metric_Coefficient(MetricParameters,Curve_Data(:,ip1)));
        NewLength = NewLength_0 + NewLength_1;

        if (NewLength < OldLength)
            Length_Reduced = 1;
            Min_Guess = Mid_Guess;
            Mid_Guess = 0.5 * (Max_Guess+Min_Guess);
            OldLength = NewLength;
            Local_Length(im1) = NewLength_0;
            Local_Length(i) = NewLength_1;
        else
            Max_Guess = Mid_Guess;
            Mid_Guess = 0.5 * (Max_Guess+Min_Guess);
        end
    end

    Number_of_Steps = Min_Guess + Number_of_Steps;

    end

    if Number_of_Steps == 0
        Curve_Data(:,i) = Curve_Data(:,i) +
            Single_Step_Direction * Delta *
            Basis_Rotation_Matrix(:,Long_Direction);
        if Single_Step_Direction > 0
            Local_Length(im1) = Fwd_NewLength(Long_Direction,1);
            Local_Length(i) = Fwd_NewLength(Long_Direction,2);
        else
            Local_Length(im1) = Bwd_NewLength(Long_Direction,1);
            Local_Length(i) = Bwd_NewLength(Long_Direction,2);
        end
        Curve_Data_Moved = 1;
    else
        Candidate_Shift(1) = 0;
        for k = 2:Dimension
            Candidate_Shift(k) = Delta * round(Shift_Gradient(k)
                * Number_of_Steps);
        end
        Candidate_Shift = Basis_Rotation_Matrix *
            Candidate_Shift;
        Curve_Data(:,i) = Curve_Data(:,i) + Candidate_Shift;
        Curve_Data_Moved = 1;
```

```
            end
        end
end
end
```

A.2 Length Function Implementation

This function returns the approximate length as described by the function L in the pseudocode. The position of the node to be moved is passed via qb, with it's left and right neighbours as qa and qc respectively. The variables $MetricCf_a, MetricCf_b$ and $MetricCf_c$ take the metric values at qa, qb and qc respectively. The variable $tmp\_energies$ returns the lengths between qa, qb and qb, qc.

```
function [left_energy right_energy] =
    L(qa,qb,qc,MetricCf_a,MetricCf_b,MetricCf_c)
     left_energy  =  0.5 * norm(qb-qa) * (MetricCf_a+MetricCf_b);
     right_energy  =  0.5 * norm(qc-qb) * (MetricCf_b+MetricCf_c);
end
```

A.3 Metric Coefficient Function for Test Cases

This is a simple implementation of the $Metric\_Coefficient$ function. The array C receives anything set in the $MetricParameters$ variable. The array x corresponds to the point at which to evaluate the metric.

```
function out = Metric_Coefficient(C,x)
        b = ones(1,C(1));
        b(1) = 0;
        b = (1/norm(b)) * b;
        out = exp(-C(2) * b * x);
end
```

A.4 Global Algorithm Implementation

This function initialises the global algorithm. The variable N specifies the number of global nodes to use. The remaining parameters behave as described earlier in this appendix.

```
function Curve_Data = Birkhoff(N, NodeOrder, Start_Point,End_Point, Const,
    Alpha, varargin)

Dimension = size(Start_Point,1);
```

```
if nargin - 6 == 0
    MetricParameters = 0;
else
    MetricParameters = varargin{1};
end

N_1 = N-1;
Tangent_Step = (End_Point - Start_Point)/(N_1);
Curve_Data = zeros(Dimension,N);
for i = 1:N
    Curve_Data(:,i) = Start_Point + (i-1) * Tangent_Step;
end

tolbar = 10^(-4);
tol = tolbar+1;
Playlist = [3:2:N-1,2:2:N-1];
OldN = 0;

while tol > tolbar
    tvect = [];
    UpdateNodes = 0;

    for i = Playlist
        [temp flag] =
            GradientJacobi(NodeOrder,Curve_Data(:,i-1),Curve_Data(:,i+1),
            Const,Alpha,MaxGradTol,MinGradTol,MetricParameters);
        tvect = [tvect norm(Curve_Data(:,i) -
            temp(:,2^(NodeOrder-1)+1),2)];
        Curve_Data(:,i) = temp(:,2^(NodeOrder-1)+1);
   end

   tol = norm(tvect,2)*OldN;
end
end
```

Appendix B

C Code Listing for Parallel Curve Shortening Algorithm

The purpose of this appendix is to give the full implementation of the parallel algorithm in C. All of the code segments should be in a single file before running.

B.1 Header Code

The following code is necessary for the implementation and defines the functions that the global algorithm will use.

```
#include <stdio.h>
#include <stdlib.h>
#include <time.h>
#include <math.h>
#include <mpi.h>

long double Metric_Coefficient(long double*C,long double*x,long int
    Dimension);

void L(long double*u,long double*v,long double*w,long double mu,long
    double mv,long double mw,long int Dimension,long double*tmp_energies);

void FindFrame(long double*Tau,long double**Basis_Rotation_Matrix,long int
    Dimension);

void Curve_Shorten(long double*Start_Point,long double*End_Point,long int
    NodeOrder,long double Constant,long double*MetricParameters,long
    double Alpha,long int Dimension, long double*Results);
```

```
int Birkhoff_Step(long double*Local_Length,long double*u,long
    double*v,long double*w,long double**Basis_Rotation_Matrix,long double
    Delta,long double Epsilon,long int Dimension,long
    double*MetricParameters,long int Current_Node,long double Constant);

void Slave_Core(int rank, long int Dimension, long double *
    MetricParameters, long double NodeOrder, long double Constant, long
    double Exponent);
```

B.2 Function to Move a Single Node and Test for Length Reduction

The following function will take a node, along with its neighbours, and attempt to find the optimal position which reduces length. The function `Birkhoff_Step` returns 1 to indicate that the length reduction was successful, otherwise it returns a 0. The position of the node to be moved is passed via v, with it's left and right neighbours as u and w respectively. The variable *Local_Length* records the optimal length's found, if any. The variable *Basis_Rotation_Matrix* represents the change of basis matrix between the standard basis of $\mathbb{R}^d$ and the frame along the initial curve. The variable *Constant* is described in the pseudocode as C. The other key variable, not seen in the pseudocode, is *MetricParameters*, which is essentially a free array which one may use to configure the metric function.

```
int Birkhoff_Step(long double*Local_Length,long double*u,long
    double*v,long double*w,long double**Basis_Rotation_Matrix,long double
    Delta,long double Epsilon,long int Dimension,long
    double*MetricParameters,long int Current_Node,long double Constant){

    long int i= 0;
    long int j= 0;

    long double**Fwd_NewLength= malloc(Dimension*sizeof(long double*));
    for(i= 0;i<Dimension;i++){
        Fwd_NewLength[i]= malloc(2*sizeof(long double));
    }
    long double**Bwd_NewLength= malloc(Dimension*sizeof(long double*));
    for(i= 0;i<Dimension;i++){
        Bwd_NewLength[i]= malloc(2*sizeof(long double));
    }
    long double*Candidate_Shift= malloc(Dimension*sizeof(long double));
    long double*Advance_Direction= malloc(Dimension*sizeof(long double));
    long double*Shift_Gradient= malloc(Dimension*sizeof(long double));
```

```
long double*Tmp_Candidate_Shift= malloc(Dimension*sizeof(long double));
long int Current_Node_m1= Current_Node-1;
long int Current_Node_p1= Current_Node+1;
long double
    OldLength=Local_Length[Current_Node_m1]+Local_Length[Current_Node];
long double NewLength;
long int Number_of_Steps= 0;
int Not_Local_Minimum= 0;
long double Forward_Length= 0;
long double Backward_Length= 0;
long double mu= Metric_Coefficient(MetricParameters,u,Dimension);
long double mw= Metric_Coefficient(MetricParameters,w,Dimension);
long double Advance_Gradient_Scaling= -((Constant*Epsilon)/(2*Delta));

Advance_Direction[0]= 0;

for(i= 1;i<Dimension;i++){
    for(j= 0;j<Dimension;j++){
            Candidate_Shift[j]= v[j]+Delta*Basis_Rotation_Matrix[j][i];
    }
    L(u,Candidate_Shift,w,mu,
        Metric_Coefficient(MetricParameters,Candidate_Shift,Dimension),
        mw,Dimension,Fwd_NewLength[i]);
    Forward_Length= Fwd_NewLength[i][0]+Fwd_NewLength[i][1];

    for(j= 0;j<Dimension;j++){
            Candidate_Shift[j]= v[j]-Delta*Basis_Rotation_Matrix[j][i];
    }
    L(u,Candidate_Shift,w,mu,
        Metric_Coefficient(MetricParameters,Candidate_Shift,Dimension),
        mw,Dimension,Bwd_NewLength[i]);
    Backward_Length= Bwd_NewLength[i][0]+Bwd_NewLength[i][1];

    if((Forward_Length<OldLength)||(Backward_Length<OldLength)){
            Not_Local_Minimum= 1;
    }

    Advance_Direction[i]=
        Advance_Gradient_Scaling*(Forward_Length-Backward_Length);

}


if(Not_Local_Minimum==1){
```

```
long int Long_Direction= 1;
long double tmp_Compare= fabsl(Advance_Direction[Long_Direction]);

for(i= 1;i<Dimension;i++){
    if(fabsl(Advance_Direction[i])> tmp_Compare){
        Long_Direction= i;
    }
}

long double Number_of_Gradient_Steps=
    ceil(fabsl(Advance_Direction[Long_Direction])/Delta);
long double Single_Step_Direction= 0;

if(Advance_Direction[Long_Direction]> 0){
    Single_Step_Direction= 1;
}else{
    Single_Step_Direction= -1;
}

long double Shift_Gradient_Scaling=
    Single_Step_Direction/Advance_Direction[Long_Direction];

for(i= 1;i<Dimension;i++){
    Shift_Gradient[i]= Shift_Gradient_Scaling*Advance_Direction[i];
}

int Length_Reduced= 1;
long double Tmp_Length[2];
long double NewLength_0;
long double NewLength_1;

while(Length_Reduced==1){
    Length_Reduced= 0;
    Number_of_Steps+= Number_of_Gradient_Steps;

    Candidate_Shift[0]= 0;
    for(i= 1;i<Dimension;i++){
        Candidate_Shift[i]=
            Delta*round((Shift_Gradient[i]*Number_of_Steps));
    }

    for(i= 0;i<Dimension;i++){
        Tmp_Candidate_Shift[i]= 0;
        for(j= 0;j<Dimension;j++){
            Tmp_Candidate_Shift[i]+=
                Basis_Rotation_Matrix[i][j]*Candidate_Shift[j];
```

```
            }
        }

        for(i= 0;i<Dimension;i++){
            Candidate_Shift[i]= Tmp_Candidate_Shift[i]+v[i];
        }

        L(u,Candidate_Shift,w,mu,
            Metric_Coefficient(MetricParameters,Candidate_Shift,
            Dimension),mw,Dimension,Tmp_Length);

        NewLength_0= Tmp_Length[0];
        NewLength_1= Tmp_Length[1];
        NewLength= NewLength_0+NewLength_1;

        if(NewLength<OldLength){
            Length_Reduced= 1;
            OldLength= NewLength;
            Local_Length[Current_Node_m1]= NewLength_0;
            Local_Length[Current_Node]= NewLength_1;
        }else{
            Number_of_Steps-= Number_of_Gradient_Steps;
        }
    }

    if(Number_of_Gradient_Steps> 1){
        long double Max_Guess=
            pow(2,(ceil(log(Number_of_Gradient_Steps)/log(2))));
        long double Min_Guess= 0;
        long double Mid_Guess= 0.5*Max_Guess;

        while((Max_Guess> Min_Guess)&&(Mid_Guess==round(Mid_Guess))){
            long double tmp_Steps= Number_of_Steps+Mid_Guess;

            Candidate_Shift[0]= 0;
            for(i= 1;i<Dimension;i++){
                Candidate_Shift[i]=
                    Delta*round((Shift_Gradient[i]*tmp_Steps));
            }

            for(i= 0;i<Dimension;i++){
                Tmp_Candidate_Shift[i]= 0;
                for(j= 0;j<Dimension;j++){
                    Tmp_Candidate_Shift[i]+=
                        Basis_Rotation_Matrix[i][j]*Candidate_Shift[j];
                }
```

```
            }

            for(i= 0;i<Dimension;i++){
                Candidate_Shift[i]= Tmp_Candidate_Shift[i]+v[i];
            }

            L(u,Candidate_Shift,w,mu,
                Metric_Coefficient(MetricParameters,Candidate_Shift,
                Dimension),mw,Dimension,Tmp_Length);

            NewLength_0= Tmp_Length[0];
            NewLength_1= Tmp_Length[1];
            NewLength= NewLength_0+NewLength_1;

            if(NewLength<OldLength){
                Min_Guess= Mid_Guess;
                Mid_Guess= 0.5*(Max_Guess+Min_Guess);
                OldLength= NewLength;
                Local_Length[Current_Node_m1]= NewLength_0;
                Local_Length[Current_Node]= NewLength_1;
            }else{
                Max_Guess= Mid_Guess;
                Mid_Guess= 0.5*(Max_Guess+Min_Guess);
            }
        }
        Number_of_Steps= Min_Guess+Number_of_Steps;
    }

    if(Number_of_Steps==0){
        long double Signed_Delta= Single_Step_Direction*Delta;
        for(j= 0;j<Dimension;j++){
            v[j]+=
                Signed_Delta*Basis_Rotation_Matrix[j][Long_Direction];
        }

        if(Single_Step_Direction> 0){
            Local_Length[Current_Node_m1]=
                Fwd_NewLength[Long_Direction][0];
            Local_Length[Current_Node]=
                Fwd_NewLength[Long_Direction][1];
        }else{
            Local_Length[Current_Node_m1]=
                Bwd_NewLength[Long_Direction][0];
            Local_Length[Current_Node]=
                Bwd_NewLength[Long_Direction][1];
        }
```

```
    for(i= 0;i<Dimension;i++){
        free(Fwd_NewLength[i]);
    }

    for(i= 0;i<Dimension;i++){
        free(Bwd_NewLength[i]);
    }

    free(Fwd_NewLength);
    free(Bwd_NewLength);
    free(Candidate_Shift);
    free(Tmp_Candidate_Shift);
    free(Advance_Direction);
    free(Shift_Gradient);

    return(1);
}else{
    Candidate_Shift[0]= 0;
    for(i= 1;i<Dimension;i++){
        Candidate_Shift[i]=
            Delta*round((Shift_Gradient[i]*Number_of_Steps));
    }

    for(i= 0;i<Dimension;i++){
        Tmp_Candidate_Shift[i]= 0;
        for(j= 0;j<Dimension;j++){
            Tmp_Candidate_Shift[i]+=
                Basis_Rotation_Matrix[i][j]*Candidate_Shift[j];
        }
    }

    for(i= 0;i<Dimension;i++){
        Candidate_Shift[i]= Tmp_Candidate_Shift[i]+v[i];
    }

    for(i= 0;i<Dimension;i++){
        v[i]= Candidate_Shift[i];
    }

    for(i= 0;i<Dimension;i++){
        free(Fwd_NewLength[i]);
    }

    for(i= 0;i<Dimension;i++){
        free(Bwd_NewLength[i]);
```

```
            }

            free(Fwd_NewLength);
            free(Bwd_NewLength);
            free(Candidate_Shift);
            free(Tmp_Candidate_Shift);
            free(Advance_Direction);
            free(Shift_Gradient);

            return(1);
        }
    }else{
        for(i= 0;i<Dimension;i++){
            free(Fwd_NewLength[i]);
        }

        for(i= 0;i<Dimension;i++){
            free(Bwd_NewLength[i]);
        }

        free(Fwd_NewLength);
        free(Bwd_NewLength);
        free(Candidate_Shift);
        free(Tmp_Candidate_Shift);
        free(Advance_Direction);
        free(Shift_Gradient);

        return(0);
    }
}
```

B.3 Local Algorithm Implementation

The purpose of this function is to perform the local algorithm. We will only discuss parameters that have not been covered in the pseudocode or previous functions in this appendix. The parameter *NodeOrder* indicates that the calculation should be performed with $2^N + 1$ nodes, where N is *NodeOrder*. The variable *Results* returns the midpoint along the local geodesic. The variable *Alpha* takes values between 2.5 and 3.

```
void Curve_Shorten(long double*Start_Point,long double*End_Point,long int
    NodeOrder,long double Constant,long double*MetricParameters,long
    double Alpha,long int Dimension,long double*Results)
{
    long int i= 0;
```

```
long int j= 0;
long int N= pow(2,NodeOrder)+1;
long int N_1= N-1;
long int N_2= N-2;
long double*Tangent_Step= malloc(Dimension*sizeof(long double));

for(i= 0;i<Dimension;i++){
    Tangent_Step[i]= (End_Point[i]-Start_Point[i])/((long double)N_1);
}
long double Epsilon= 0;
for(i= 0;i<Dimension;i++){
    Epsilon+= Tangent_Step[i]*Tangent_Step[i];
}
Epsilon= sqrt(Epsilon);
long double Delta= pow(Epsilon,Alpha);

long double**Curve_Data= malloc(N*sizeof(long double*));
for(i= 0;i<N;i++){
    Curve_Data[i]= malloc(Dimension*sizeof(long double));
}
for(i= 0;i<N;i++){
    for(j= 0;j<Dimension;j++){
        Curve_Data[i][j]= Start_Point[j]+((long
            double)i)*Tangent_Step[j];
    }
}
long double**Basis_Rotation_Matrix= malloc(Dimension*sizeof(long
    double*));
for(i= 0;i<Dimension;i++){
    Basis_Rotation_Matrix[i]= malloc(Dimension*sizeof(long double));
}

FindFrame(Tangent_Step,Basis_Rotation_Matrix,Dimension);

long double Total_Length;
long double*Local_Length= malloc(N_1*sizeof(long double));
long int ip1= 1;
long int ip2= 2;

for(i= 0;i<N_2;i+= 2){
    L(Curve_Data[i],Curve_Data[ip1],Curve_Data[ip2],
        Metric_Coefficient(MetricParameters,Curve_Data[i],Dimension),
        Metric_Coefficient(MetricParameters,Curve_Data[ip1],Dimension),
        Metric_Coefficient(MetricParameters,Curve_Data[ip2],Dimension),
        Dimension,&Local_Length[i]);
    ip1+= 2;
```

```
    ip2+= 2;
}

int Curve_Data_Moved= 1;
long int Current_Node= 1;
long int Current_Node_m1;
long int Current_Node_p1;
long int Play_List[N_2];
long int Position= 0;

for(i= 1;i<N_1;i+= 2){
    Play_List[Position]= i;
    Position++;
}
for(i= 2;i<N_1;i+= 2){
    Play_List[Position]= i;
    Position++;
}
long int count = 0;

while(Curve_Data_Moved==1 && count < pow(N,2)){
    Curve_Data_Moved= 0;
    count++;
    for(i= 0;i<N_2;i++){
        Current_Node= Play_List[i];
        if(Birkhoff_Step(Local_Length,Curve_Data[Current_Node-1],
            Curve_Data[Current_Node],Curve_Data[Current_Node+1],
            Basis_Rotation_Matrix,Delta,Epsilon,Dimension,
            MetricParameters,Current_Node,Constant)==1){
            Curve_Data_Moved= 1;
        }
   }
}

for(i= 0;i<N;i++){
    Total_Length+= Local_Length[i];
}

N = pow(2,NodeOrder-1);

for (i = 0; i < Dimension; i++) {
    Results[i] = Curve_Data[N][i];
}

for(i= 0;i<N;i++){
    free(Curve_Data[i]);
```

```
    }

    free(Curve_Data);
    free(Local_Length);
    free(Tangent_Step);
    for(i= 0;i<Dimension;i++){
        free(Basis_Rotation_Matrix[i]);
    }
    free(Basis_Rotation_Matrix);
}
```

B.4 Metric Coefficient Function for Test Cases

This is a simple implementation of the $Metric\_Coefficient$ function. The array C receives anything set in the $MetricParameters$ variable. The array x corresponds to the point at which to evaluate the metric.

```
long double Metric_Coefficient(long double*C,long double*x,long int
    Dimension)
{
    long int i= 0;
    long double metric_value= 0;
    long double alpha= C[0];
    long double m= 1/sqrt((long double)Dimension-1);
    long double n[Dimension];

    n[0]= 0;
    or(i= 1;i<Dimension;i++){
        n[i]= m;
    }
    for(i= 0;i<Dimension;i++){
        metric_value+= x[i]*n[i];
    }

    return(exp(-alpha*metric_value));
}
```

B.5 Length Function Implementation

This function returns the approximate length as described by the function L in the pseudocode. We will only discuss parameters that have not been covered in the pseudocode or previous functions in this appendix. The variables mu, mv and mw take the metric

values at u, v and w respectively. The variable *tmp_energies* returns the lengths between u, v and v, w.

```
void L(long double*u,long double*v,long double*w,long double mu,long
    double mv,long double mw,long int Dimension,long double*tmp_energies)
{
    long int i= 0;
    long double tmp_left[Dimension];
    long double tmp_right[Dimension];
    long double left_norm= 0;
    long double right_norm= 0;

    for(i= 0;i<Dimension;i++){
        tmp_left[i]= u[i]-v[i];
        tmp_right[i]= v[i]-w[i];
    }

    for(i= 0;i<Dimension;i++){
        left_norm+= tmp_left[i]*tmp_left[i];
        right_norm+= tmp_right[i]*tmp_right[i];
    }

    left_norm= sqrt(left_norm);
    right_norm= sqrt(right_norm);
    tmp_energies[0]= 0.5*left_norm*(mu+mv);
    tmp_energies[1]= 0.5*right_norm*(mv+mw);
}
```

B.6 Function to Find Orthonormal Frame Along Initial Curve

This function takes an array Tau and returns a matrix $Basis\_Rotation\_Matrix$ whose columns form an orthonormal basis of $\mathbb{R}^{Dimension}$. In particular, this orthonormal basis contains the normalised form of Tau.

```
void FindFrame(long double*Tau,long double**Basis_Rotation_Matrix,long int
    Dimension)
{
    long int i= 0;
    long int j= 0;
    long int k= 0;
    long int l= 0;

    while(Tau[i]==0){
```

```
    i++;
}

long double A[Dimension][Dimension];
for(l= 0;l<Dimension;l++){
    A[l][j]= Tau[l];
}
for(j= 1;j<Dimension;j++){
    if(k==i){
        k++;
    }
    for(l= 0;l<Dimension;l++){
        if(l==k){
            A[l][j]= 1;
        }else{
            A[l][j]= 0;
        }
    }
    k++;
}

long double R[Dimension][Dimension];
long double v[Dimension];
long double norm_v;
for(i= 0;i<Dimension;i++){
    for(j= 0;j<Dimension;j++){
        Basis_Rotation_Matrix[i][j]= 0;
        R[i][j]= 0;
    }
    v[i]= 0;
}

for(j= 0;j<Dimension;j++){
    for(i= 0;i<Dimension;i++){
        v[i]= A[i][j];
    }
    for(i= 0;i<j;i++){
        for(k= 0;k<Dimension;k++){
            R[i][j]= R[i][j]+Basis_Rotation_Matrix[k][i]*A[k][j];
        }
        for(k= 0;k<Dimension;k++){
            v[k]= v[k]-R[i][j]*Basis_Rotation_Matrix[k][i];
        }
    }
    norm_v= 0;
    for(i= 0;i<Dimension;i++){
```

```
            norm_v+= v[i]*v[i];
         }
        norm_v= sqrt(norm_v);
        R[j][j]= norm_v;
        for(i= 0;i<Dimension;i++){
            Basis_Rotation_Matrix[i][j]= (1/norm_v)*v[i];
        }
    }
}
```

B.7 Global Algorithm Implementation - Local Algorithm on Slave Cores

The purpose of this implementation is to allow all bar one core to run the local algorithm. In particular it handles the communication between the cores and calling the `Curve_Shorten` function. All variables are described in previous sections of this appendix.

```
void Slave_Core(int rank, long int Dimension, long double *
    MetricParameters, long double NodeOrder, long double Constant, long
    double Alpha)
{
    int tag;
    long int i = 0;
    long int j = 0;
    long double EndPoints[2*Dimension+1];
    long double Mid_Point[Dimension+2];
    long double tmp_Val[Dimension];
    long double Left_Point[Dimension];
    long double Right_Point[Dimension];
    long double NodeNumber = 0;

    MPI_Status status;
    MPI_Recv(EndPoints, 2*Dimension+1, MPI_LONG_DOUBLE, 0, MPI_ANY_TAG,
        MPI_COMM_WORLD,&status);
    tag = status.MPI_TAG;

    NodeNumber = EndPoints[2*Dimension];

    while (NodeNumber > 0) {
        for (i = 0; i < Dimension; i++){
            Left_Point[i] = EndPoints[i];
        }
        for (i = Dimension; i < 2*Dimension; i++){
```

```
            Right_Point[i-Dimension] = EndPoints[i];
        }

        Curve_Shorten(Left_Point,Right_Point,NodeOrder,Constant,
            MetricParameters,Alpha,Dimension,tmp_Val);

        for (i = 0; i < Dimension; i++) {
            Mid_Point[i] = tmp_Val[i];
        }
        Mid_Point[Dimension] = NodeNumber;
        Mid_Point[Dimension+1] = tmp_Val[Dimension];

        MPI_Send(Mid_Point, Dimension+2, MPI_LONG_DOUBLE, 0, tag,
            MPI_COMM_WORLD);
        MPI_Recv(EndPoints, 2*Dimension+1, MPI_LONG_DOUBLE, 0,
            MPI_ANY_TAG, MPI_COMM_WORLD,&status);
        tag = status.MPI_TAG;

        NodeNumber = EndPoints[2*Dimension];
    }
}
```

B.8 Global Algorithm Implementation - Code for Master Core

This section of code is responsible for initialising the global algorithm. It also provides the time elapsed and error between the final curve and the example problem. The variable D sets the dimension of the problem. The variable exponent takes a value between 2.5 and 3. The variable TOL_MAX determines the stop point for the code.

```
int main(int argc, char*argv[])
{

    int np, rank;
    MPI_Init(&argc, &argv);
    MPI_Comm_rank(MPI_COMM_WORLD, &rank);
    MPI_Comm_size(MPI_COMM_WORLD, &np);

    long double * MetricParameters = malloc(1 * sizeof(long double));

    MetricParameters[0] =0.65;

    long int D = 3;
```

```
long int NodeOrder = 6;
long double Constant = 4;
long double Exponent = 2.6;
long int N = pow(2,3)+1;
long int MIN_N = (np-1)*2+2;

 long double TOL_MAX = pow(10,-3);

if (N < MIN_N) {
    if (rank == 0) {
        printf("Error: Too few nodes. Either reduce number of cores,
            or increase number of nodes.\n");
    }
    MPI_Finalize();
    return 0;
}

if (rank == 0){
    clock_t start= clock();
    long double err= 0;
    long int i = 0;
    long int j = 0;
    long int k = 0;
    long double * Start_Point = malloc(D * sizeof(long double));
    long double * End_Point = malloc(D * sizeof(long double));
    long double * Tangent_Step = malloc(D * sizeof(long double));
    long double * Mid_Node = malloc(D * sizeof(long double));
    long double ** Curve_Data = malloc(N * sizeof(long double *));
    long int NumberOfEvenNodes, NumberOfOddNodes;

    if (N%2 == 0) {
        NumberOfEvenNodes = (N-2)/2;
        NumberOfOddNodes = (N-2)/2;
    } else {
        NumberOfEvenNodes = (N-3)/2;
        NumberOfOddNodes = (N-3)/2+1;
    }

    for (i = 0; i < N; i++){
       Curve_Data[i] = malloc(D * sizeof(long double));
    }

    long double N_1 = 1/((long double)N-1);

    // Initialise Start_Point and End_Point - Replace according to
        problem
```

```
Start_Point[0] = -1;
End_Point[0] = 1;
for (i = 1; i < D; i++){
    Start_Point[i] = 0;
    End_Point[i] = 0;
}

for (i = 0; i < D; i++){
    Tangent_Step[i] = N_1*(End_Point[i]-Start_Point[i]);
}
for(i= 0;i<N;i++){
    for(j= 0;j<D;j++){
        Curve_Data[i][j]= Start_Point[j]+((long
            double)i)*Tangent_Step[j];
    }
}

long double TOL_VECT_NORM_LOC = 0;
long double TOL_VECT_NORM_GLO = TOL_MAX + 1;
int temp, tag, who;
MPI_Status status;
long int tmp_NodeNumber;
long double * tmp_InVector = malloc((D+2) * sizeof(long double));
long double * tmp_OutVector = malloc((2*D+1) * sizeof(long
    double));
long double GradientFlag;
long int * Play_List_Even = malloc(NumberOfEvenNodes*sizeof(long
    int));
long int * Play_List_Odd = malloc(NumberOfOddNodes*sizeof(long
    int));
long int Position= 0;

for(i= 2;i<N-1;i+= 2){
    Play_List_Even[Position]= i;
    Position++;
}
Position= 0;
for(i= 1;i<N-1;i+= 2){
    Play_List_Odd[Position]= i;
    Position++;
}

long double * FlaggedNodes = malloc(N * sizeof(long double));
long int ptr_shift = 0;
long double ** tmp_CD;
```

```
while(TOL_VECT_NORM_GLO > TOL_MAX){

    TOL_VECT_NORM_GLO = 0;
    TOL_VECT_NORM_LOC = 0;

    for (i=0; i<np-1; i++) {
        tmp_NodeNumber = Play_List_Even[i];

        for (j = 0; j < D; j++) {
            tmp_OutVector[j] = Curve_Data[tmp_NodeNumber-1][j];
        }
        for (j = D; j < 2*D; j++) {
            tmp_OutVector[j] = Curve_Data[tmp_NodeNumber+1][j-D];
        }
        tmp_OutVector[2*D] = tmp_NodeNumber;

        MPI_Send(tmp_OutVector, 2*D+1, MPI_LONG_DOUBLE, i+1, i,
            MPI_COMM_WORLD);
    }

    while (i<NumberOfEvenNodes) {

        MPI_Recv(tmp_InVector, D+2, MPI_LONG_DOUBLE,
            MPI_ANY_SOURCE, MPI_ANY_TAG,MPI_COMM_WORLD, &status);

        who = status.MPI_SOURCE;
        tag = status.MPI_TAG;

        tmp_NodeNumber = (long int) tmp_InVector[D];
        for (j = 0; j < D; j++){
            TOL_VECT_NORM_LOC += pow(Curve_Data[tmp_NodeNumber][j]
                - tmp_InVector[j],2);
            Curve_Data[tmp_NodeNumber][j] = tmp_InVector[j];
        }

        TOL_VECT_NORM_GLO += TOL_VECT_NORM_LOC;
        TOL_VECT_NORM_LOC = 0;

       tmp_NodeNumber = Play_List_Even[i];

        for (j = 0; j < D; j++) {
            tmp_OutVector[j] = Curve_Data[tmp_NodeNumber-1][j];
        }
        for (j = D; j < 2*D; j++) {
            tmp_OutVector[j] = Curve_Data[tmp_NodeNumber+1][j-D];
```

```
        }
        tmp_OutVector[2*D] = tmp_NodeNumber;

        MPI_Send(tmp_OutVector, 2*D+1, MPI_LONG_DOUBLE, who, i+1,
            MPI_COMM_WORLD);

        i++;
    }

    for (i=1; i<np; i++) {
        MPI_Recv(tmp_InVector, D+2, MPI_LONG_DOUBLE, i,
            MPI_ANY_TAG,MPI_COMM_WORLD, &status);
        who = status.MPI_SOURCE;
        tag = status.MPI_TAG;

        tmp_NodeNumber = tmp_InVector[D];

        for (j = 0; j < D; j++){
            TOL_VECT_NORM_LOC += pow(Curve_Data[tmp_NodeNumber][j]
                - tmp_InVector[j],2);
            Curve_Data[tmp_NodeNumber][j] = tmp_InVector[j];
        }

        TOL_VECT_NORM_GLO += TOL_VECT_NORM_LOC;
        TOL_VECT_NORM_LOC = 0;
    }

    for (i=0; i<np-1; i++) {

        tmp_NodeNumber = Play_List_Odd[i];

        for (j = 0; j < D; j++) {
            tmp_OutVector[j] = Curve_Data[tmp_NodeNumber-1][j];
        }
        for (j = D; j < 2*D; j++) {
            tmp_OutVector[j] = Curve_Data[tmp_NodeNumber+1][j-D];
        }
        tmp_OutVector[2*D] = tmp_NodeNumber;

        MPI_Send(tmp_OutVector, 2*D+1, MPI_LONG_DOUBLE, i+1, i,
            MPI_COMM_WORLD);
    }

    while (i<NumberOfOddNodes) {
```

```
        MPI_Recv(tmp_InVector, D+2, MPI_LONG_DOUBLE,
            MPI_ANY_SOURCE, MPI_ANY_TAG,MPI_COMM_WORLD, &status);

        who = status.MPI_SOURCE;
        tag = status.MPI_TAG;

        tmp_NodeNumber = (long int) tmp_InVector[D];

        for (j = 0; j < D; j++){
            TOL_VECT_NORM_LOC += pow(Curve_Data[tmp_NodeNumber][j]
                - tmp_InVector[j],2);
            Curve_Data[tmp_NodeNumber][j] = tmp_InVector[j];
        }

        TOL_VECT_NORM_GLO += TOL_VECT_NORM_LOC;
        TOL_VECT_NORM_LOC = 0;

        tmp_NodeNumber = Play_List_Odd[i];

       for (j = 0; j < D; j++) {
            tmp_OutVector[j] = Curve_Data[tmp_NodeNumber-1][j];
        }
        for (j = D; j < 2*D; j++) {
            tmp_OutVector[j] = Curve_Data[tmp_NodeNumber+1][j-D];
        }
        tmp_OutVector[2*D] = tmp_NodeNumber;

        MPI_Send(tmp_OutVector, 2*D+1, MPI_LONG_DOUBLE, who, i+1,
            MPI_COMM_WORLD);

        i++;
    }

    for (i=1; i<np; i++) {
        MPI_Recv(tmp_InVector, D+2, MPI_LONG_DOUBLE, i,
            MPI_ANY_TAG,MPI_COMM_WORLD, &status);
        who = status.MPI_SOURCE;
        tag = status.MPI_TAG;

        tmp_NodeNumber = tmp_InVector[D];

        for (j = 0; j < D; j++){

            TOL_VECT_NORM_LOC += pow(Curve_Data[tmp_NodeNumber][j]
                - tmp_InVector[j],2);
```

```
                Curve_Data[tmp_NodeNumber][j] = tmp_InVector[j];

            }

            TOL_VECT_NORM_GLO += TOL_VECT_NORM_LOC;
            TOL_VECT_NORM_LOC = 0;
        }


        TOL_VECT_NORM_GLO = sqrt(TOL_VECT_NORM_GLO)*N;
}

 tmp_OutVector[2*D] = -1;
 for (i=1; i<np; i++) {
        MPI_Send(&tmp_OutVector, 2*D+1, MPI_LONG_DOUBLE, who,
            N+1,MPI_COMM_WORLD);
 }

 // Compute time-elasped and error for test problem - Remove for
        other problems

 printf("Time elapsed:
     %f\n",((double)clock()-start)/CLOCKS_PER_SEC);
 for(i= 0;i<N;i++){
        err+=
        fabs(Curve_Data[i][1]-(1/(MetricParameters[0]*sqrtl(D-1)))*
            log((cos(MetricParameters[0]*Curve_Data[i][0]))/
            (cos(MetricParameters[0]))));
 }
 err= (1/(long double)N)*err;
 printf("Average Error: %Lf\n\n",err);

 free(Start_Point);
 free(End_Point);
 free(Tangent_Step);
 free(Mid_Node);
 free(tmp_OutVector);
 free(tmp_InVector);
 free(Play_List_Even);
 free(Play_List_Odd);
 free(FlaggedNodes);
 free(tmp_CD);
 free(MetricParameters);
 for (i = 0; i < N; i++){
        free(Curve_Data[i]);
 }
```

```
        free(Curve_Data);
    } else {
        Slave_Core(rank,D,MetricParameters,NodeOrder,Constant,Exponent);
    }

    return(0);
}
```

Bibliography

[ACM09] M. Amar, G. Crasta, and A. Malusa. On the Finsler metric obtained as limits of chessboard structures. *Adv. Calc. Var.*, 2:321–360, 2009.

[All92] G. Allaire. Homogenization and two-scale convergence. *SIAM J. Math. Anal.*, 23:1482–1518, 1992.

[Apo90] T. Apostol. *Modular functions and Dirichlet series in number theory.* Springer-Verlag, 1990.

[Arn97] V. I. Arnold. *Mathematical methods of classical mechanics.* Springer-Verlag, 1997.

[AV98] M. Amar and E. Vitali. Homogenization of periodic Finsler metrics. *J. Convex Anal.*, 5(1):171–186, 1998.

[BA90] A. Banerjee and N.P. Adams. Dynamics of classical systems based on the principle of stationary action. *J. Chem. Phys.*, 92(7330), 1990.

[BBF02] Andrea Braides, Giuseppe Buttazzo, and Ilaria Fragalá. Riemannian approximation of Finsler metrics. *Asymptotic Analysis*, 31:177–187, 2002.

[BBI01] Dmitri Burago, Yuri Burago, and Sergi Ivanov. *A course in metric geometry.* American Mathematical Society, 2001.

[BCS00a] A.M. Balk, A.V. Cherkaev, and L.I. Slepyan. Dynamics of chains with non-monotone stress-strain relations. I. Model and numerical experiments. *J. Mech. Phys. Solids*, 49:131–148, 2000.

[BCS00b] A.M. Balk, A.V. Cherkaev, and L.I. Slepyan. Dynamics of chains with non-monotone stress-strain relations. II. Nonlinear waves and waves of phase transition. *J. Mech. Phys. Solids*, 49:149–171, 2000.

[BD78] Giuseppe Buttazzo and G. Dal Maso. Γ-limit of a sequence of non-convex and non-equi-Lipschitz integral functionals. *Ric. Mat.*, 27:235–251, 1978.

[BD98] Andrea Braides and Anneliese Defranceschi. *Homogenisation of multiple integrals*. Oxford University Press, 1998.

[BDF91] E. Barvínek, I. Daler, and Jan Franců. Convergence of sequences of inverse functions. *Archivum Mathematicum*, 27:201–204, 1991.

[BGLS06] J.-F. Bonnans, J.C. Gilbert, C. Lemarechal, and C.A. Sagastizàbal. *Numerical optimization*. Number XIV in Universitext. Springer-Verlag, 2006.

[Bir66] George D. Birkhoff. *Dynamical Systems*, volume IX of *American Mathematical Society Colloquium Publications*. American Mathematical Society, 1966.

[BPF01] Giuseppe Buttazzo, Luigi De Pascale, and Ilaria Fragalá. Topological equivalence of some variational problems involving distances. *Discrete and Continuous Dynamical Systems*, 7(2):247–258, April 2001.

[Bra86] Andrea Braides. A homogenisation theorem for weakly almost periodic functionals. *Rend. Accad. Naz. Sci. XL*, 104:261–281, 1986.

[Bra92] Andrea Braides. Almost periodic methods in the theory of homogenization. *Appl. Anal.*, 47:259–277, 1992.

[Bra02] Andrea Braides. Γ-*convergence for beginners*. Oxford University Press, 2002.

[CB03] Bogdan Craciun and Kaushik Bhattacharya. Homogenisation of a Hamilton-Jacobi equation associated with the geometric motion of an interface. *Proc. Roy. Soc. Edinburgh Sect. A*, 133A:773–805, 2003.

[Dal93] G. Dal Maso. *An introduction to* $\Gamma-$*convergence.* Birkhäuser, 1993.

[Dav04] Andrea Davini. *Finsler metrics in Optimisation Problems and Hamilton-Jacobi equations*. PhD thesis, University of Pisa, 2004.

[DF75] E. DeGiorgi and T. Franzoni. Su un tipo di convergenza variazionale. *Atti Accad. Naz. Lincei Rend. Cl. Sci.*, 58:842–850, 1975.

[DP07] A. Davini and M. Ponsiglione. Homogenization of two-phase metrics and applications. *Journal d'Analyse Mathématique*, 103(1):157–196, 2007.

[DS73] G. Dal Maso and S. Spagnolo. Sulla convergenza degli integrali dell'energia per operatori ellittici del secondo ordine. *Boll. Un. Mat. Ital.*, 8:391–411, 1973.

[E91] Weinan E. A class of homogenisation problems in the calculus of variations. *Comm. Pure Appl. Math.*, XLIV:733–759, 1991.

[Eva89] Lawrence Craig Evans. The perturbed test function method for viscosity solutions of nonlinear PDE. *Proc. Roy. Soc. Edinburgh Sect. A*, 111:359–375, 1989.

[Eva92] Lawrence Craig Evans. Periodic homogenisation of certain fully nonlinear PDE. *Proc. Roy. Soc. Edinburgh Sect. A*, 120:245–265, 1992.

[Eva98] Lawrence Craig Evans. *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society, 1998.

[Fis97] A. Fischer. *Die hybride Monte-Carlo-Methode in der Molekülphysik*. PhD thesis, Freie Universität Berlin, 1997.

[Fol84] G.B. Folland. *Real Analysis: Modern Techniques and their Applications*. Wiley Interscience, 1984.

[FS40] F. Förster and E. Scheil. Untersuchung des zeitlichen Ablaufes von Umklappvorgängen in Metallen. *Zeitschrift für Metallkunde*, 6:165–173, 1940.

[FS02] D. Frenkel and B. Smit. *Understanding Molecular Simulation*, volume 1. Academic Press, Orlando, FL, USA, 2002.

[GE01] Diogo Gomes and Lawrence Craig Evans. Effective Hamiltonians and averaging for Hamiltonian dynamics I. *Arch. Ration. Mech. Anal.*, 157(1):1–33, 2001.

[GE02] Diogo Gomes and Lawrence Craig Evans. Effective Hamiltonians and averaging for Hamiltonian dynamics II. *Arch. Ration. Mech. Anal.*, 161(4):271–305, 2002.

[GO04] Diogo Gomes and Adam Oberman. Computing the effective Hamiltonian: A varational approach to homogenization. *SIAM J. Control Optim.*, 43(3):792–812, 2004.

[Gom00] Diogo Gomes. *Viscosity solutions and asymptotics of Hamiltonian systems*. PhD thesis, University of California, 2000.

[GW92] R.E. Gillian and K.R. Wilson. Shadowing, rare events, and rubber bands. A variational Verlet algorithm for molecular dynamics. *J. Chem. Phys.*, 97(1757), 1992.

[HMSZ13] M. Herrmann, K. Matthies, H. Schwetlick, and J. Zimmer. Subsonic phase transition waves in bistable lattice models with small spinodal region. Submitted, 2013.

[Jos05] J. Jost. *Riemannian Geometry and Geometric Analysis*. Universitext. Springer-Verlag, fourth edition, 2005.

[KP02] S.G. Krantz and H.R. Parks. *A primer of real analytic functions*. Birkhäuser Boston Inc., 2002.

[LPV88] Pierre-Louis Lions, G Papanicolaou, and S R S Varadhan. Homogenisation of Hamilton-Jacobi equations. Preprint, 1988.

[MR99] J. E. Marsden and T. S. Ratiu. *Introduction to mechanics and symmetry: a basic exposition of classical mechanical systems*. Springer-Verlag, 1999.

[Mur78] F. Murat. Compacité par compensation. *Ann. Scuola Norm. Sup. Pisa. Cl. Sci.*, 5:489–507, 1978.

[OTV09] Adam Oberman, Ryo Takei, and Alaxander Vladimirsky. Homogenisation of metric Hamilton-Jacobi equations. *Multiscale Model. Simul.*, 8(2):269–295, 2009.

[PK84] M. Peyrard and M. Kruskal. Kink dynamics in the highly discrete sine-gordon system. *Physica D*, 14:88–102, 1984.

[SCC05] L.I. Slepyan, A.V. Cherkaev, and E. Cherkaev. Transition waves in bistable structures. II. Analytical solution: wave speed and energy dissipation. *J. Mech. Phys. Solids*, 53:407–436, 2005.

[Sle02] L.I. Slepyan. *Models and phenomena in fracture mechanics*. Springer-Verlag, 2002.

[Spa68] S. Spagnolo. Sulla convergenza delle soluzioni di equazioni paraboliche ed ellittiche. *Ann. Scuola Norm. Sup. Pisa. Cl. Sci.*, 22:571–597, 1968.

[SSZ12] H. Schwetlick, D.C. Sutton, and J. Zimmer. Nonexistence of heteroclinic travelling waves for a bistable Hamiltonian lattice model. *J. Nonlinear Sci.*, 22:917–934, 2012.

[SZa] H. Schwetlick and J. Zimmer. The computation of long time Hamiltonian trajectories for molecular systems via global geodesics. Numerical Mathematics and Advanced Applications 2011 (ENUMATH, Leicester), to appear.

[SZb] H. Schwetlick and J. Zimmer. A convergent string method: Existence and approximation for the Hamiltonian boundary-value problem. Preprint.

[SZ07] H. Schwetlick and J. Zimmer. Solitary waves for nonconvex FPU lattices. *J. Nonlinear Sci.*, 17:1–12, 2007.

[SZ09a] H. Schwetlick and J. Zimmer. Calculation of long time classical trajectories: Algorithmic treatment and applications for molecular systems. *J. Chem. Phys.*, 130(124104), 2009.

[SZ09b] H. Schwetlick and J. Zimmer. Existence of dynamic phase transitions in a one-dimensional lattice model with piecewise quadratic interaction potential. *SIAM J. Math. Anal.*, 41:1231–1271, 2009.

[SZ12] H. Schwetlick and J. Zimmer. Kinetic relations for a lattice model of phase transitions. *Arch. Ration. Mech. Anal.*, 206(2):707–724, 2012.

[Tar79] L. Tartar. *Compensated compactness and applications to partial differential equations*, volume 39 of *Res. Notes in Math.*, pages 136–211. Pitman, London, 1979.

[Tar90] L. Tartar. H-measures, a new approach for studying homogenisation, oscillations and concentration effects in partial differential equations. *Proc. Roy. Soc. Edinburgh Sect. A*, 115:193–230, 1990.

[TV05] L. Truskinovsky and A. Vainchtein. Kinetics of martensitic phase transitions: Lattice model. *SIAM J. Appl. Math.*, 66:533–553, 2005.

[Vai10] A. Vainchtein. The role of the spinodal region in the kinetics of lattice phase transitions. *J. Mech. Phys. Solids*, 58:227–240, 2010.

[Ven91] S. Venturini. Derivation of distance functions in $\mathbb{R}^N$. Preprint, 1991.

[VHRT98] A. Vainchtein, T. Healy, P. Rosakis, and L. Truskinovsky. The role of the spinodal region in one-dimensional martensitic phase transitions. *Physica D*, 115:29–48, 1998.

[VK12] A. Vainchtein and P.G. Kevrekidis. Dynamics of phase transitions in a piecewise linear diatomic chain. *J. Nonlinear Sci.*, 22:107–134, 2012.

[Zem65] A.H. Zemanian. *Distribution theory and transform analysis.* McGraw-Hill, 1965.